AN ARCHITECTURE MODEL OF THE U.S. AIR TRANSPORTATION NETWORK

A Dissertation
Presented to
The Academic Faculty

by

Kisun Song

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Daniel Guggenheim School of Aerospace Engineering

Georgia Institute of Technology
May 2020

AN ARCHITECTURE MODEL OF THE U.S. AIR TRANSPORTATION NETWORK

Approved by:

Prof. Dimitri Mavris, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Holger Pfaender
School of Aerospace Engineering
Georgia Institute of Technology

Prof. Daniel Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Eric Upton
Preliminary Design
Gulfstream Aerospace

Dr. Jung-Ho Lewe
School of Aerospace Engineering
Georgia Institute of Technology

Date Approved: November 15th, 2019

"It would be very remarkable if any system existing in the real world could be exactly represented by any simple model. However, cunningly chosen parsimonious models often do provide remarkably useful approximations."

– George Box –

To my father, my mother, and my sister,

I love you and thank you.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Professor Dimitri Mavris for his dedicated advice, support, and guidance. Without his guidance and persistent help including finance, this dissertation would not have been possible. I would like to also express my appreciation to the committee members, Professor Daniel Schrage, Dr. Jung-Ho Lewe, Dr. Holger Pfaender, and Dr. Eric Upton. Their comments and advice have significantly helped me finish my dissertation.

I would like to express my special gratitude to Dr. Jung-Ho Lewe. Being advised by him for the entire period of my Ph.D. program was one of the greatest happiness of my life. He provided lots of opportunities for me to contribute not only to the research but also to the funded project, the smart campus. Also, I would be grateful for Dr. Scott Duncan and all colleagues in the smart campus project. Dr. Scott Duncan took care of me in successfully working on the funded project.

My acknowledgement also goes to my Korean fellows in ASDL, Jung-Hyun Kim and Seul-Ki Kim for their support. As the Ph.D. students in the ASDL, we were such a great team not only to perform our research but also to enjoy our lives together in many ways. Also, I would like to thank my dear friend, Bobae Kang, for his support. I always enjoyed working, studying, and talking with him on any topic. I hope we keep it going in the future.

Last but not least, my final acknowledgement goes to my beloved family which is the most important reason in my life. Without their endless support and love, I could not

have finished my Ph.D. program. I love all my family members: my father Euisup Song, my mother Syung-Soon In, my sister Yinae Song, my brother-in-law Joon-Ho Kim, and my nephews Suah Kim and Min-Hoo Kim. I love you all.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AA	American Airlines
ACAIS	Air Carrier Activity Information System
ADA	Airline Deregulation Act
AFD	Airport Facilities Directory
AIP	Airport Improvement Program
ANN	Artificial Neural Network
AS	Alaska Airlines
ATN	Air Transportation Network of the United States of America
B6	JetBlue Airways
BTS	Bureau of Transportation Statistics
CA	Continental Airlines
CAA	Civil Aeronautics Acts
CONUS	CONtiguous U.S. territory
CPD	Cross-track Projection Distance
DB1B	Airline Origin and Destination Survey
DL	Delta Airlines
DOT	Department of Transportation
ERA	Environmentally Responsible Aviation
F9	Frontier Airlines

FAA	Federal Aviation Administration
FL	AirTran Airways
G4	Allegiant Air
GA	Genetic Algorithm
H&S	Hub-and-Spoke
HA	Hawaiian Airlines
HWB	Hybrid Wing Body
LCC	Low-Cost-Carriers
M&A	Mergers and Acquisitions
ML	Machine Learning
NAH	New Aviation Horizon
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NK	Spirit Airlines
NW	Northwest Airlines
O-D	Origin-and-Destination
OAI	Office of Airline Information
P2P	Point-to-Point
QueSST	Quiet Supersonic Technology
R&D	Research & Development
RE	Research Experiment
RH	Research Hypothesis
ROC	Receiver Operating Characteristic
RQ	Research Question
SDB1B	Symmetric DB1B

SST	SuperSonic Transport
T-100D	Domestic Air Carrier Statistics (Form 41 Traffic)
T-100I	International Air Carrier Statistics (Form 41 Traffic)
TRL	Technology Readiness Level
TWA	Trans World Airlines
UA	United Airlines
US	United States of America
US	US Airways
V&V	Validation & Verification
VX	Virgin America
WN	SouthWest Airlines

Greek Symbols

γ	Disutility discernibility controller
λ	Demand in DB1B
υ	Network policy controller
$\hat{\sigma}$	Great circle distance
τ	Demand
$\tau^{[P]}$	Primary demand
$\tau^{[S]}$	Secondary demand
τ^{SST}	Demand for supersonic air travel
Ψ	Set of aircraft types
ψ	Element of set of aircraft types
φ	Airport hub discount factor

Roman Symbols

$\mathscr{A}$	Binary adjacency
$\mathscr{A}^w$	Weighted adjacency
$\mathscr{k}$	Node degree
$\mathscr{s}$	Strength
$\mathscr{a}^w$	Entry of weighted adjacency matrix
$\mathscr{c}$	Binary clustering coefficient
$\mathscr{c}^w$	Weighted clustering coefficient
$\mathscr{C}_B$	Betweenness centrality
$\mathscr{C}_C$	Closeness centrality
$\mathscr{E}$	Enplanement (volume)
$\mathscr{B}$	Segment bandwidth
$\mathscr{G}$	Gravity of nodes
lon	Longitude
lat	Latitude
α	Airport
α_{Acc}	Access airport
$\mathscr{E}^{SST}$	Enplanement for supersonic air travel
G	Generated demand/volume
r	Route
s	Segment
N	Number of airports
L	Number of segments
PDF	Probability distribution function

CDF	Cumulative density function
DU	Route disutility
F	Route fitness
ENV	Evolutionary environment
ATN	Simulated air transportation network
A	Airport matrix
T	Demand matrix
t	Evolution time step
D	Great circle distance
f	Final time step
C	Connection volume
$Cost$	Aircraft hourly cost
Cap	Aircraft capacity
Rng	Aircraft range
Spd	Aircraft cruise speed
Yr	Aircraft debut year
Blt	Aircraft block time
POL	Airlines' network construction policy
ATN_R	Reference air transportation network
ATN_S	Simulated air transportation network
ft	Aircraft segmental flight time
$AlnCost$	Airline-side disutility
LF	Aircraft load factor
$PaxCost$	Passenger-side disutility
ToF	Average time for transfer

VoT	Monetary value of time
SVF	Service viability factor
A_{ϕ}	Major hub airports where hub discount factor is applied
err_i	Individual relative error
err_T	Global relative error
$path$	Supersonic shortest path sequence
ft^{SST}	Flight time of supersonic transport
A^{SST}	Viable airports for supersonic ATN

SUMMARY

For almost a century, the U.S. Air Transportation Network (ATN) has continuously and successfully adapted to its changing environment as if it were a living organism. Today, the complexity of the network encompasses various exogenous as well as endogenous factors: fuel price, socioeconomic and political climates, atmospheric conditions, varying interests of stakeholders, and growing dependence on technology, to name a few. Its sophisticated interactions among diverse factors affecting the ATN have captivated many network researchers. Some researchers have attempted to retrieve an order out of seemingly chaotic constructions, while others have analyzed historical variations in its properties to understand the ATN's behavioral mechanisms.

However, its mathematical representation led by the known components and rules is yet to be developed. Thus, this thesis develops an architecture model of the ATN that mathematically represents the components and rules with realism. In the model, the network evolves in a virtual environment comprising three time-variant components – demand, airport, and aircraft technology – built upon extensive realistic datasets. Then the network is constructed by the active agents – airlines – performing multi-tiered network evolutionary processes and evolves into a strong hub-and-spoke (H&S) structure network that mimics the function of its reference: real-world ATN. The validated model provides various opportunities to conduct extensive analyses and studies on the past, current, and future of the ATN.

Finally, a case study has been performed: forecasting the future ATN disruption caused by the technological revolution of civil supersonic transports. It provided an

opportunity to experience the exploratory and interpretative capability of the architecture model, which shed light on performing future researches with better realism.

CHAPTER 1. INTRODUCTION

1.1 Air Transportation Network Overview

1.1.1 Air Transportation Network as a Complex System

The U.S. air transportation network (ATN) is a complex system in transportation networks, where the nodes and the segments are the airports and the actual directed flight routes between airports, respectively. Its complexity can also be easily identified by the fact that the ATN exhibits the two prominent properties in a complex network: small-world and scale-free properties. [1,2,3] The complexity stems from various endogenous factors such as diverse stakeholders with distinct roles and objectives, high dependency on technology, and the importance of seasonality on schedule as well as exogenous ones, including fuel price, socioeconomic/political settings, and weather.

The ATN is also represented by its Hub-and-Spoke (H&S) structure that evolved over history; the centralization of the resources into a small number of top major airports and major airlines handle most transportation demand. Figure 1 shows a chart about the airline domestic market share between August 2018 and July 2019. [4] For readers' understanding, key terms of the two distinguished structures of the ATN – H&S and Point-to-Point (P2P) – are compared in Table 1, which were adopted from reference. [5]

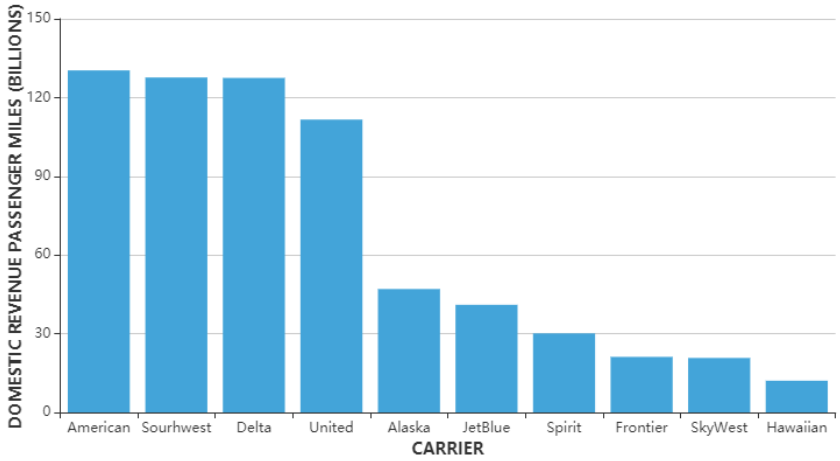


Figure 1 – Airline domestic market share (August 2018 ~ July 2019.)

Table 1 – Characteristics of Hub-and-Spoke and Point-to-Point route systems.

Attribute	Hub and Spoke	Point to Point
Scope	Optimized by connecting service to a wide geographical area and many destinations	Each route serves a single city pair. Individual routes may be dispersed.
Connectivity	Most passengers connect at the hub(s) for a continuing flight(s) to destination	No connections provided (although incidental or "rolling hub" connections are standard)
Dependence	Each route highly dependent on other routes for connecting passengers	Routes operate independently, and traffic is not affected by demand from other routes
Demand	Varying demand in any given city pair may be offset by demand from other markets	Only varying frequency and pricing available to counter demand variance
Market Size	Efficiently serves cities of greatly varying size	Requires high-density markets with at least one endpoint being a high demand origin/destination
Frequency	Supports high daily frequency to all destinations	Generally lower frequency depending on the market type and density
Pricing	Frequency and coverage appeal to business travelers providing a margin for higher business fares	Both business and leisure passengers are generally price- seeking
Asset Utilization	Limited by network geography, connection timing, and hub congestion	No network constraints on utilization
Fleet Operation	Large range in seating capacity is necessary to match capacity with traffic, usually requires more than one fleet type	Suited to a single fleet type

The complexity of the strong H&S structure can also be characterized by the competitions of different airlines that have different markets, different aircraft fleet, different network construction policies, and different hub locations (e.g., operating bases.) For instance, the domestic-flight passengers for Southwest Airlines and American Airlines in 2018 were 97.3% and 80.4%, respectively. [6] Delta Airlines and Spirit Airlines operate 18 and 3 different aircraft types, respectively. [7] United Airlines and Hawaiian Airlines operate in 194 and 27 U.S. airports, respectively. [7,8] Moreover, different deployments of the evolution of them in terms of network topology even escalates the level of complexity of the ATN much more, as Figure 2 shows some examples. [9,10,11,12,13,14]

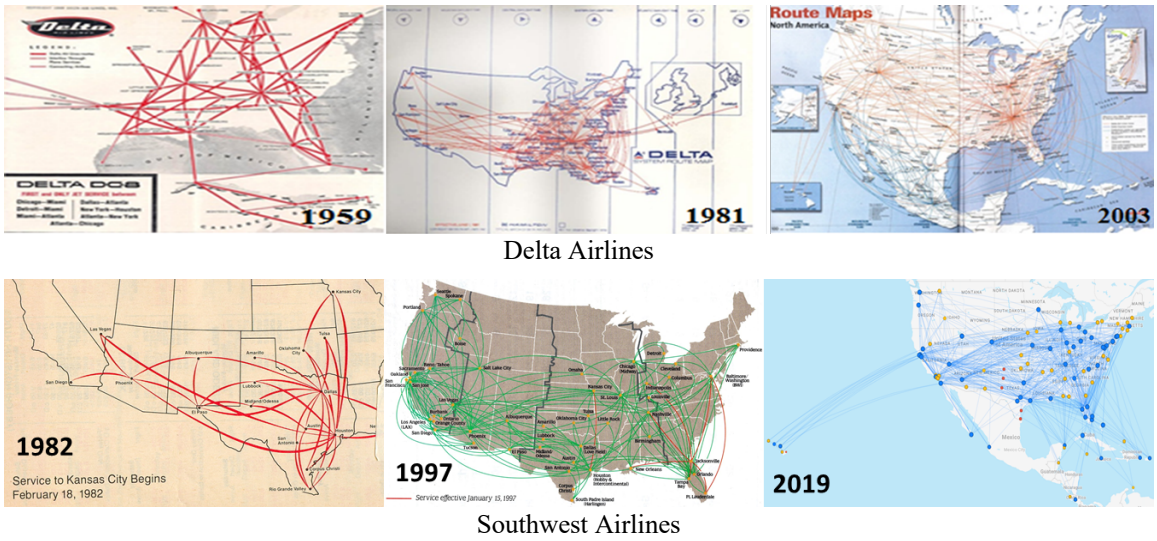


Figure 2 – Evolutions of different U.S. carriers.

1.1.2 Importance of ATN: A Look from Statistical Perspectives

ATN is one of the largest spatial networks in the modern era. From an economic perspective, it is a critical enabler to achieve economic growth and development of a society. It is the fastest and unique inter-continental enabler for integrating the global

economy and providing strong connections between different places in an international as well as a domestic scale. According to the statistical data by the Bureau of Transportation Statistics (BTS), the U.S. scheduled passenger airlines reported a first-quarter 2019 net profit of $3.3 billion and a net operating revenue of $44.4 billion. [15] Its yearly sum is 0.93% of the Gross Domestic Product (GDP) of the U.S. Moreover, as of August 2019, 21 U.S. air carriers have 742,431 employees. [16] Despite a considerable amount of disruption caused by the 9/11 terror attack, the recession in 2008, its economic status has rebounded in 2010 and exceeded afterward.

In a topological perspective, the Federal Aviation Administration (FAA) announces that there are over 44,000 daily flights are performed, five thousand aircraft are in the sky at peak operational times, 2.78 million daily passengers fly in and out of U.S. airports, 518 airport traffic control towers work, 154 terminal radar approach control facilities work, 5.3 million square miles of U.S. domestic airspace is covered. [17] All these enormous amounts of network flow are accommodated by 7,397 commercial aircraft and 212,875 general aviation fleet, which are operated by 633,317 active pilots certified by the FAA. [18]

1.2 Research Motivation

1.2.1 Estimating Entry-Level Metrics for Long-Term and Short-Term Policy

The FAA forecasts that the total volume (enplanements) of the domestic network will reach 1.12 trillion by the year 2039, as shown in Figure 3. [19] Hence, airports could confront many challenges like capacity issue and delay problem so that stakeholders are required to make policies to prepare for future aptly.

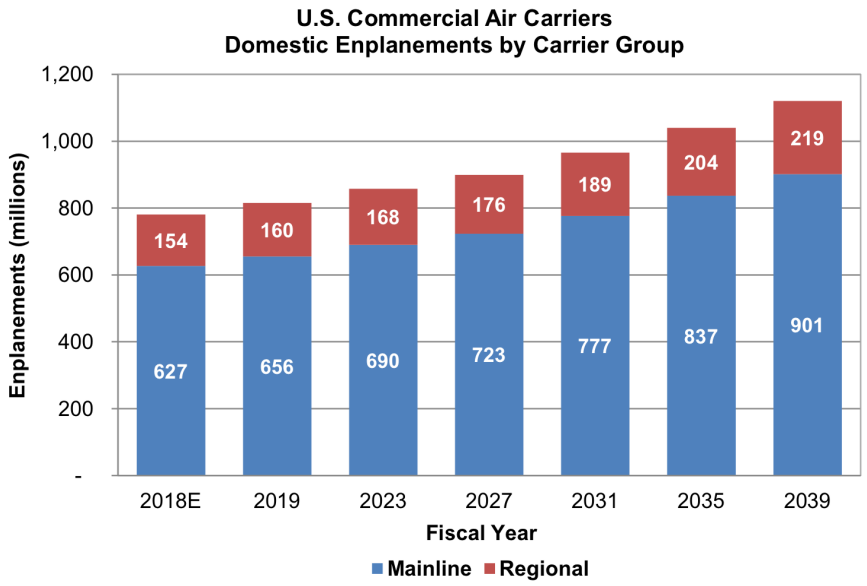


Figure 3 – Forecast of the U.S. domestic civil aviation enplanements until 2039.

The United Kingdom (U.K.) government conducted a consultation for making long-term and short-term policies to prepare for the future growth of U.K. aviation. The report expects the total volume of the U.K. ATN to reach 435 million in the year 2050 with the runway expansion in the LHR airport, as shown in Figure 4. [20]

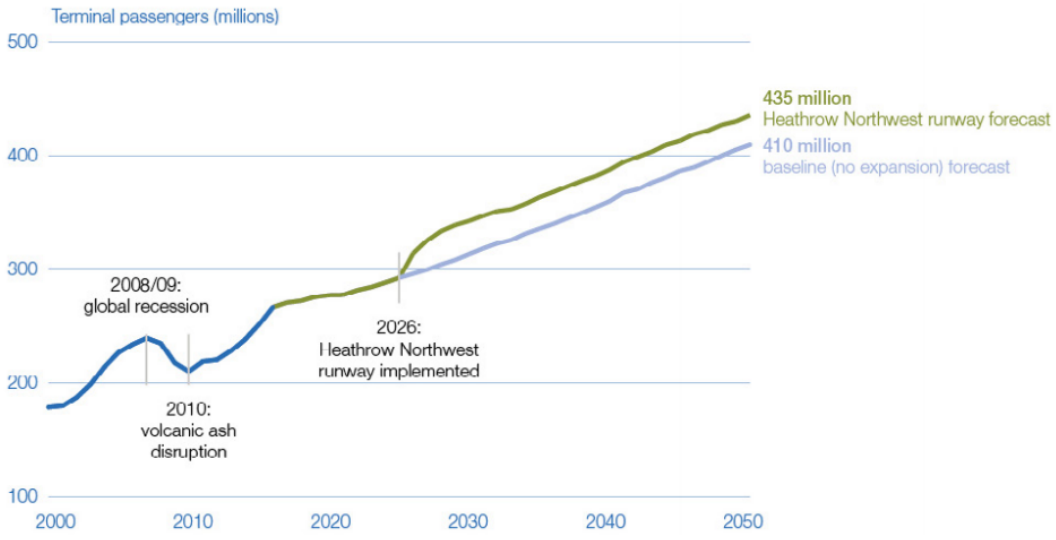


Figure 4 – Trends in volume at U.K. airports and forecast to 2050.

Meanwhile, the National Aeronautics and Space Administration (NASA) has proposed the Environmentally Responsible Aviation (ERA) project, which aims to explore the feasibility, benefits, and risks of future vehicle technologies to have the potential to alleviate the environmental impact. The technological readiness level of this project is tabulated in Figure 5 with the subsonic transport system-level metric goals. [21,22] All of the proposed goals will lead to an array of introductions of technologically advanced aircraft into the ATN market.

TECHNOLOGY BENEFITS	TECHNOLOGY GENERATIONS (Technology Readiness Level = 5-6)		
	Near-term 2015-2025	Mid-term 2025-2035	Far-term Beyond 2035
Noise (cumulative below Stage 4)	22 - 32 dB	32 - 42 dB	42 - 52 dB
LTO NO_x Emissions (below CAEP 6)	70 - 75%	80%	>80%
Cruise NO_x Emissions (relative to 2005 best in class)	65 - 70%	80%	>80%
Aircraft Fuel/Energy Consumption (relative to 2005 best in class)	40 - 50%	50 - 60%	60 - 80%

Figure 5 – NASA's subsonic transport system-level metrics.

The enplanement forecasts by the FAA and the U.K. government employ the system-level volume to make/adjust their future air transportation network policies. As to NASA's ERA project, the technological emission goals are also described by system-level metrics. Although these metrics are surely informative, the expected enplanement by the forecasts cannot tell about the segmental enplanements, and the technological emission goals of the ERA project hardly encourage to estimate the environmental impact on specific terrains of the U.S. continent.

1.2.2 Estimating the Impact of Advanced Aircraft on Future ATN

The New Aviation Horizon (NAH) initiative of the NASA [23, 24] is a project that seeks to validate the aircraft technological innovations for reducing fuel use, emissions, and noise in aircraft design and operation. It provides a blueprint where details of the innovative technologies envisioned until the year 2026, as illustrated in Figure 6.

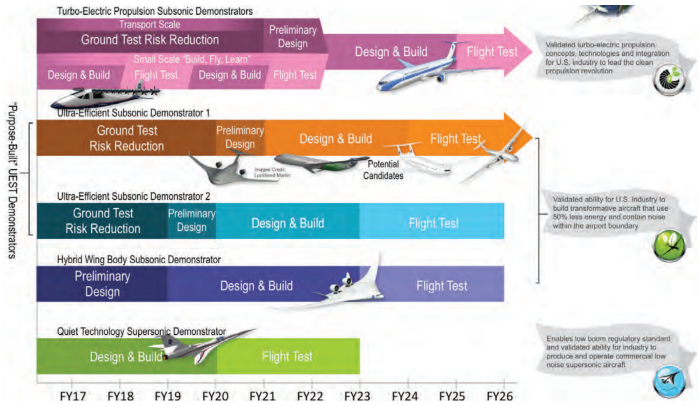


Figure 6 – NASA's plan for QUESST development.

STARC

Ultra-Efficient Hybrid Wing Body (HWB)

Quiet Supersonic Technology (QueSST)

SCEPTOR

Figure 7 – Four focused aircraft designed in NASA's NAH initiative.

Also, the leading next-generation aircraft types in NAH are seen in Figure 7. Each concept embraces different innovative technologies. For example, harnessing superconductors is actively researched to apply as one of the promising technology candidates for propulsion in the next-generation aircraft. [25]

As such, the manufacturers must reach the target TRL on time. However, no engineering environment can explore and forecast the network disruption in response to the introduction of those technologies into the ATN market: competition against the established dominant sub-sonic aircraft, topological variation, demise/rise of airports, new premium markets enabled by enhanced performance. Therefore, the capability to perform the market analysis by the aircraft being designed in the future should be one of the most significant factors in aircraft R&D, as notionally illustrated in Figure 8. An example comparison advocates this statement: the history of yearly deliveries of two state-of-the-art commercial aircraft types, Boeing B787 Dreamliner and Airbus A380, as shown in Table 2. [26,27]

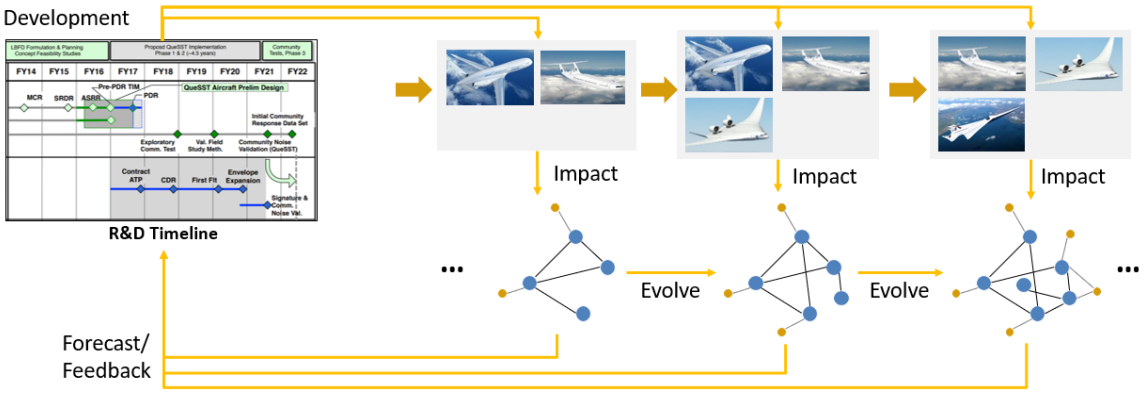


Figure 8 – Notional processes of designing future aircraft considering the technological impact on the airline network market.

Table 2 – Yearly deliveries of B787 and A380.

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
B787	-	-	-	-	3	46	65	114	135	137	136	145
A380	1	12	10	18	26	30	25	30	27	28	15	12

1.3 Research Objective

The research motivations arise from different perspectives and goals, but they commonly emphasize the need to have a model that represents the ATN in the real world. However, mathematically modeling the full granularity of the structure of the complex ATN into its every detail is not realistic or maybe even impossible. Thus, it would be more realizable not only to abstract the fundamental components and rules properly but also to represent their critical interactions that dominate the deployment of ATN evolution. Therefore, the research objective of this thesis is **to develop an architecture model that can mathematically represent the evolution of the U.S. air transportation network led by the components and rules with realism.** Finally, this thesis will provide its following key research questions and hypotheses in section 2.4.

1.4 Research Contributions

A list of research contributions is summarized as the following tuples. Note that detail descriptions are provided in CHAPTER 7. This thesis:

- Develops an architecture model that mathematically represents the fundamental components and rules of the U.S. air transportation network with a small number of design parameters rather than thousands of modeling parameters.

- Introduces a multi-tiered network evolution approach that decomposes the ATN into its sub-network tiers to formulate the mathematical rules more relevantly.
- Develops a simulation-based ATN design and analysis framework based on the collaboration between the Javascript programming language and functional programming philosophy.
- Enhances realism of the evolutionary information of airports and demand – compared to the established researches – by abstracting the real history into the architecture model.
- Introduces a concept of evolution of aircraft (e.g., historical technological advancement) in the architecture model that consecutively updates the critical specifications along time.
- Considers airport capacity constraint into the architecture model, led to an associated experiment.
- Enhances the realism by being able to tackle any changes of the complex demand history (increase/decrease.)
- Performs a reliable forecast about the future deployment of the evolution of the ATN that could be disrupted by supersonic transports with its augmented realism.

1.5 Dissertation Structure

The remainder of this thesis is organized as follows; CHAPTER 2 conducts a literature review on various perspectives of the ATN, and CHAPTER 2 interconnects the observations made from the literature review with the gaps and formulates the primary

research statements of this thesis: research questions and hypotheses. CHAPTER 3 introduces the details of the fundamental components of the architecture model. CHAPTER 4 explains the main procedural mechanisms of the developed architecture model for incorporating the components and the mathematical rules altogether. CHAPTER 5 simulates the actual evolution of the network from 1917 to 2018, validate and verify the model, and performs several experiments to prove the research hypotheses, eventually. CHAPTER 6 conducts a simulation to forecast the future of the network disruption by supersonic transports in a simplified manner. Finally, CHAPTER 7 summarizes the overall conclusions of the research of this thesis and provides the academic contributions and recommended future works.

CHAPTER 2. LITERATURE REVIEW

The literature review is performed in a macro-scale perspective; micro-scale topics such as delay propagation and crew scheduling are not considered since they are beyond the scope of this research. As an introduction, the following important observations are made:

1. Researching the complex ATN can be represented as studying the evolution of the strong H&S structure in terms of the components as well as the rules.
2. Decomposing the ATN into multiple tiers can help understand its complex properties as it allows researchers to focus on specific characteristics.
3. The ATN evolution has been influenced by a multitude of internal and external events while preserving the mainstream of the growth of the H&S structure.
4. ATN logistics optimization research enabled complex H&S-structured ATN topology to be designed by numerically solving complex optimization problems under the given circumstances.
5. ATN topology design research focused on formulating the mechanisms of creating network segments by harnessing the known knowledge from analyzing the reference ATN.

2.1 ATN Structural Analysis

2.1.1 Characterizing the H&S Structure

2.1.1.1 ATN as Complex Network

The scale-free property was firstly proposed by Barabasi and Albert [1] while Watts and Strogatz developed the small-world property. [28] A network is considered scale-free if the distribution of degree or the probability for a randomly selected node to have a degree of k is approximately following a power rule described in the following equation:

$$p(k) = ak^{-\gamma} \tag{2.1}$$

Here, p is the probability distribution, a is the normalization constant, and $2 \leq \gamma \leq 3$ is the parameter that characterizes the scale-free property. The power rule explicitly describes that a small number of nodes in the network have many neighbors, and the degree significantly drops in most nodes, eventually making the distribution follow the power rule.

As to the Watts-Strogatz model, the generated small-world property is distinct from the randomly generated network, typically the one by Erdos and Renyi. [29] A network is considered having 'small-world' property when most nodes are not directly connected, but most nodes can be traversed to every other by passing through a small number of common ones. In many cases, this property is identified by juxtaposing the relationship between the clustering coefficient and the average shortest path length of every node with that of a random network generated by the Erdos-Renyi model that has the same number of segments and nodes. The small-world property shows high clustering and low average shortest path and commonly observed in real-world networks such as the power grid and biological neural networks.

Many studies confirmed that the ATN shows small-world and scale-free characteristics found in other networks. [28,30,31,32,33] These studies used various

metrics to analyze the structure of the network in topological and functional perspectives. The former perspective considers only the network components, nodes, and segments, whereas the latter considers the weight or importance of them together in research, whose metrics are mostly based on the established fundamental research about complex networks. [1,34]

Guimera et al. [35] investigate the global structure of the worldwide air transportation network. Starting from discovering the scale-free small-world network characteristics from the worldwide air transportation network, the paper discovers some anomalous characteristics by analyzing betweenness centrality. Anomalous characteristics mean that the nodes with more connections are not always the most central in the network. The paper conjectures a multi-community structure of the network as the root of this behavior and then argues that the multi-community structure in the ATN cannot be fully explained solely based on geographical constraints and that geopolitical considerations must be addressed.

2.1.1.2 Evolution of H&S Structure

In addition to observing the small-world and scale-free characteristics, there are many more elusive characteristics of the complex ATN, often represented by the strong H&S structure. As mentioned, the overall characteristics of the Hub-and-Spoke (H&S) structure of the ATN have been analyzed the most. The explicit implementation of well-formulated network metrics and properties has enhanced our level of understanding of the ATN. Various centrality metrics such as betweenness and closeness were used to examine the

evolutionary property of the ATN compared to other types of networks, thereby explaining how the strong H&S structure has matured over history. [33,36]

Billie and Kincaid [37] analyze the U.S. ATN mainly based on the historical changes. The paper revisits the emergence and advancement of the U.S. airline industry while summarizing historical events such as the outbreak of World War I, deregulation, and their consequences. The authors discuss how the airline industry under the criteria of complex networks.

Cheung and Gunes [38] analyze the U.S. ATN by exploring the statistical data for three years: 1991, 2001, and 2011. Using various metrics to investigate the historical changes, this paper identifies the growth of the ATN by embracing the increase of the number of airports and flight segments. Similarly, Gegov et al. [39] analyze the ATN by exploring its trend through the years of 1990, 2000, and 2010 using network metrics. Metrics of interest to view the growth of networks include the number of airports, total connections, connected airports, average degree, average hops, clustering, the number of airports without connections, and scaling factors. The metrics are populated in multiple two-dimensional charts to help one visually grasp the trends of the ATN. One crucial observation is that the ranked passenger distribution appears to follow a logarithmic trend, implying high heterogeneity in passengers on different connections.

Bonnefoy and Hansman [30] focus on the capacity issue as a critical point in the air transportation system. The capacity issue is of the concern that the ATN will hardly be able to meet the forecasted future demand. The paper investigates the airports' scaling mechanisms and the factors that influence the structure of the network. The investigation

shows that airports have been able to accommodate the demand by increasing capacity and improving operation efficiency while satisfying the infrastructure constraints. The authors suggest a possible solution from the perspective of viewing the ATN as an aggregated multi-airport level system.

Studies [40,41,42] show that the topological complexity of the current ATN has remained stable after the restructuring, which occurred in around 2002. In other words, the number of airports and segments – links, arcs, connections – is stationary. The current ATN confronts various issues mostly associated with its function: demand, enplanement, cost, operation, airport capacity. These studies capture the impact on the network from the complex, convoluted interaction between those factors, which reinforce the centrality of the ATN by juxtaposing the topological metrics with functional ones; the former is not sensitive to history (i.e., small-world and scale-free properties), whereas the latter changes over time.

Recently, researchers are using Machine Learning (ML) techniques to analyze the vast and complex H&S networks. Guimera and Sales-Pardo [43] propose a mathematical and computational framework to deal with the problem of the reliability of source data in complex networks. The authors demonstrate that an algorithm can identify both 'missing' and 'spurious' interactions between nodes in noisy networks.

OBSERVATION 1: Researching the complex ATN can be represented as studying the evolution of the strong H&S structure in terms of the components as well as the rules.

2.1.2 Multiple Sub-Networks of the ATN

The overwhelmingly complex structure of the ATN encouraged researchers to embrace a new approach: decompose the network into multiple sub-networks. Many researchers attempted to consider the ATN as a multi-tiered network, not a single-tiered one. Here, the 'multiple tiers' are not an empirical finding but rather a conceptual approach to understand the ATN's dynamics and topology. [44,45] For this reason, the criteria for decomposing tiers can be subjective.

Sales-Pardo et al. [46] introduce an unsupervised ML method to extract the hierarchical organization of the complex systems, one of which is the world-wide ATN. By validating the proposed ML method with the pre-established ensemble of network hierarchically nested random graphs for validation, the authors demonstrate that the developed network analysis approach identifies the hierarchical structure of the complex ATN eclipsed with an acceptable level of accuracy.

Lordan and Sallan [47] apply the *k*-core decomposition algorithm [48] to decompose the ATN into the multiple sub-networks and to identify the core cities. They also perform a robustness analysis to find critical cities. After the decomposition, the differences between the subnetworks are analyzed based on geographical and socio-economic factors. Finally, the authors discuss how the network robustness under the multi-level structure can be improved. Figure 9 shows the core, bridge, and periphery of the ATN of North America. The figure is adopted from reference. [47]

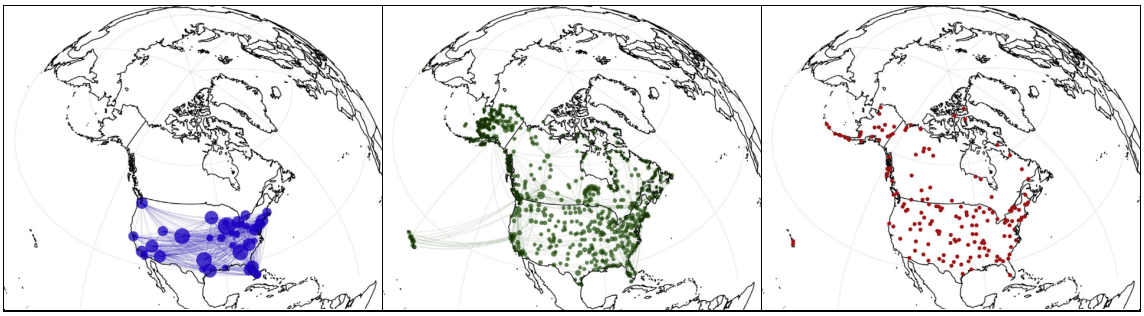


Figure 9 – Core, bridge, and periphery networks of North America.

Verma et al. [49] reveal the redundant and resilient network structure of the world-wide ATN by identifying the significant long-distance travels between the core cities and the insignificant short-distance travels between regional cities that heavily rely on the well-established connectivity of the core. The authors employ the *k*-core decomposition algorithm is employed for extracting the sub-networks of the world ATN to its three network tiers: core, bridge, and periphery. The authors confirm that the core sub-network contains only 2.3% of the total number of all airports in the world but covers the majority of the entire network volume. Figure 10 shows the decomposed network structure. The figure is adopted from reference. [49]

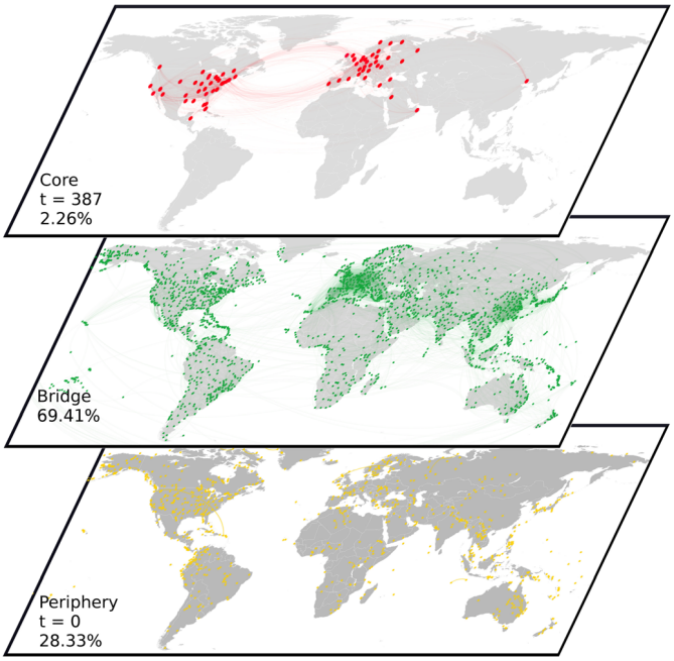
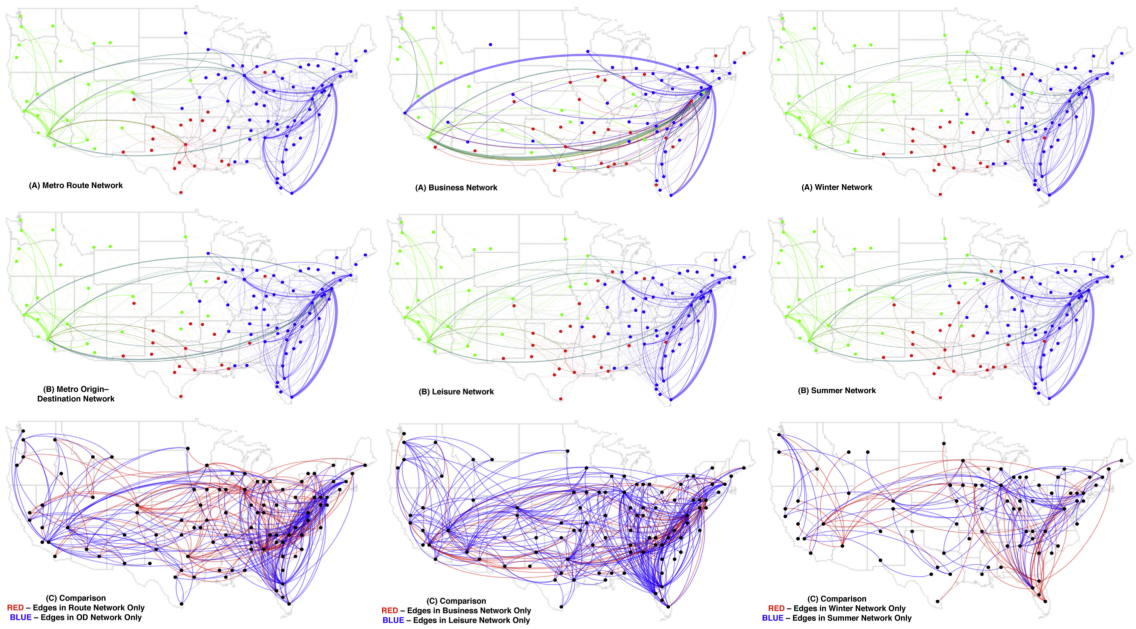


Figure 10 – The world airline network decomposed into three sub-networks.

Neal [33] considers the ATN as a collection of different types of air transportation systems based on the scale (airport vs. metropolitan area), species (business vs. leisure), and season (summer vs. winter). Motivated by the limitation that the past research has focused on the network of routes flown between airports, the author attempts to analyze the differences and similarities of the various aspects of the ATN topology. Claiming that the similarities mask different characteristics that sophisticated network metrics indicate, the author argues that differences among the various ATN topologies should be understood in detail. Figure 11 shows the comparison of different network tiers of the ATN.



Direct vs. indirect Business vs. leisure Winter vs. summer

Figure 11 – Various Subnetworks in the U.S. ATN.

No matter how the multi-level structure of the ATN is determined, the goal is to make an overwhelmingly complex entity into a collection of its simpler ones. Different sub-network tiers can be explained and studied by simpler hypotheses or approaches. Likewise, even though the ATN is a single entity of which the components and rules change over time, the multi-tier network decomposition approach is reasonable to tackle the complexity of the ATN more effectively.

OBSERVATION 2: Decomposing the ATN into multiple tiers can help understand its complex properties as it allows researchers to focus on specific characteristics.

2.1.3 ATN Disruption due to Phenomenal Events

Another important research topic is the change of the ATN after critical events, which are categorized into exogenous socio-economic events and endogenous technological paradigm shifts.

2.1.3.1 Exogenous: Socio-Economic Events

As the first historical event in the U.S. ATN evolution, the Air Mail Act of 1925, also known as the Kelly Act, was a piece of critical legislation that has dramatically incubated the momentum of civil air transportation network. It was dedicated to the airmail service rather than to civil aviation. This legislation has opened an era where private airlines made contracts with the national post service department to provide the scheduled flight operations for the first time in history. [50] These contracts not only demonstrated the overall safety of scheduled flights but also facilitated the U.S. government to establish the infrastructures – civil airways, navigational aids, regulations – for the development of civil aviation network. [51,52,53] As a result, the Civil Aeronautics Acts (CAA) has been enacted in 1926. In 1958, Congress enacted the Federal Aviation Act, which was nearly identical to the CAA, with an exception that it is more focused on the coverage of safety.

The integrated historical data can be found from 1978 when the Airline Deregulation Act (ADA) has been enacted as a part of a regulatory reform movement, including the banking, telecommunications, and energy industries. The Act led to many industry consolidations and the accelerated structural shift of the ATN from P2P to H&S. [50,54]

Goetz and Sutton [55] investigated the structural change of the ATN before and after the ADA by comparing the volumetric network change in various metrics. The paper

demonstrates that major domestic hub cities (e.g., Dallas, Chicago, and Atlanta) and international gateway cities (e.g., Los Angeles, New York City, and San Francisco) have emerged as a result of the airline company consolidations as well as geographical rearrangement of the H&S structure that followed the ADA. These cities became the core centers of the ATN, with spoke cities in the vicinity. The relative advantage of the hubs and the gateway airports included increased air transportation employment, frequency of service, passenger flow, and lower fares. Figure 12 shows the termination of civil flight services, and Figure 13 shows the absolute change in flight departures as a consequence of what the ADA has brought about.

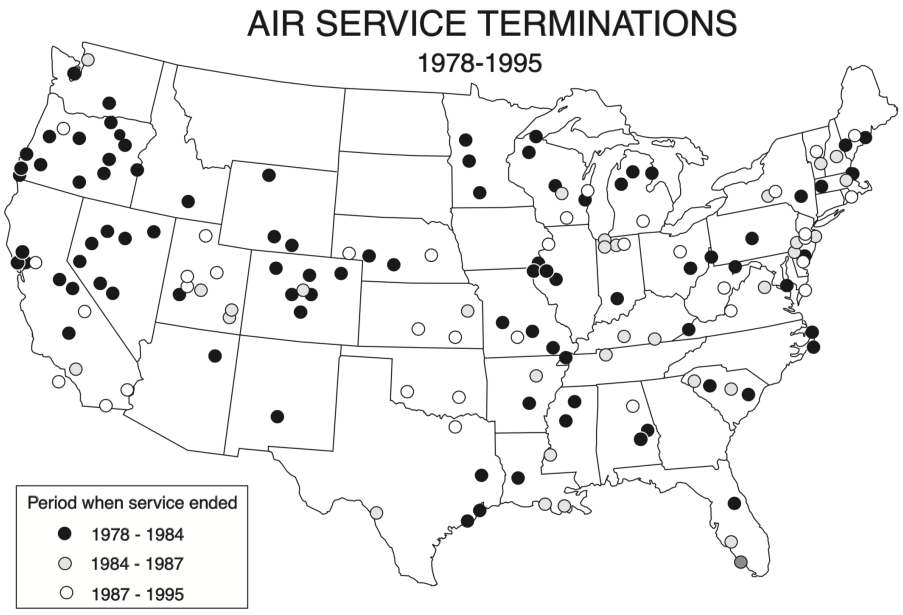


Figure 12 – Passenger-airline service terminations from 1978–1995.

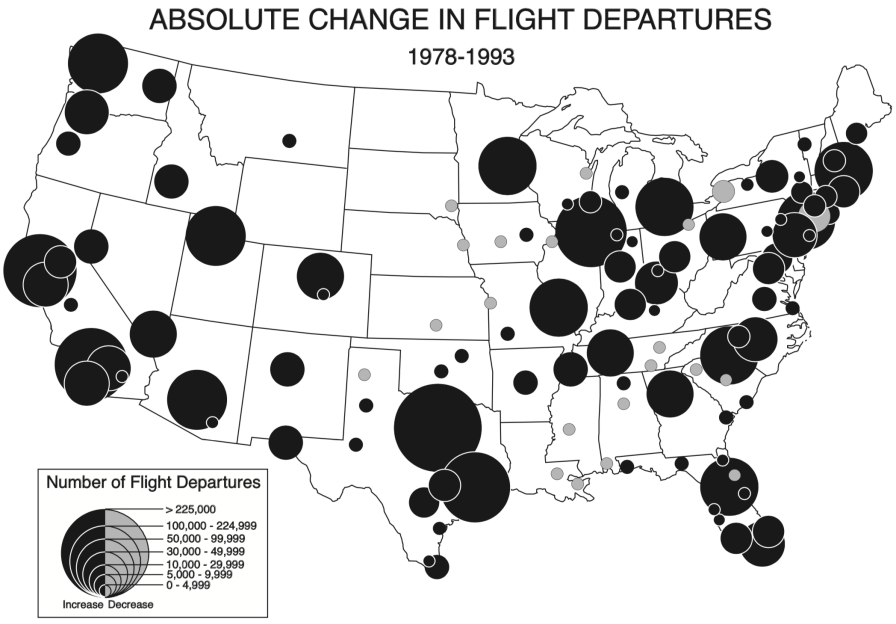
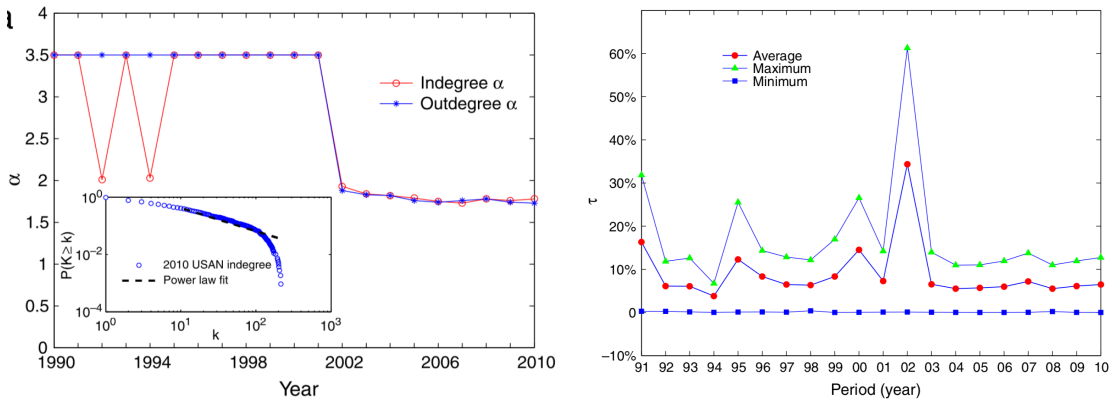


Figure 13 – Absolute changes in flight departures from 1978–1993 for the 114 largest air-passenger cities.

Different metrics have identified this massive shift. Reynolds-Feighan [56] investigated the centralization of network volume by correlating the Gini index of centralization and the network volume distribution among airlines. The Gini index of centralization has increased from an average of 0.48 in 1969 for large hubs to an average of 0.85 in 1993. This indicates that the high degree of concentration has further intensified after the ADA, by centralizing most of the air travel demand into the large hubs such as Atlanta and Chicago.

Another significant socio-economic event in the history of the ATN is the 9/11 tragedy in 2001, which caused an unprecedented disruption on the ATN. [57,58,59] Studies confirm the distinct change of network restructuring in 2001. Compared the ATN in 2001, the ATN in 2002 showed a significant drop in the average flight distance and a steep increment of the number of airports and segments, enhancing the network resiliency and

robustness to exogenous attack. Meanwhile, network efficiency has been significantly abated. Figure 14 illustrates the discernible changes. Figures were migrated from reference. [57]



History of scale-free property

History of structural change values

Figure 14 – Distinct change in 2001 ~ 2002 captured by metrics.

2.1.3.2 Endogenous: Evolution of Components and Rules

As to the components that are highly correlated with the function of the ATN, analyzing the variation of components of the ATN is also a relevant research field for understanding the underlying dynamics of the ATN. For example, ATN has relied on the advancement of aerospace technology, which adopted by airlines. [60,61] Brueckner [62] studies the impact on the fleet routing of airlines by regional jet, a new technological innovation in the ATN evolution. The author discovers evidence that regional jets have enabled passengers to benefit from better service quality via higher flight frequencies. Regional jets in the main H&S structure of the ATN were so attractive that they could replace the legacy turboprop aircraft in the regional markets.

Azzam [63] focuses on studying the different roles of airports in the global mainstream of the ATN evolution by using a 29-year record of data. The calculated conditional transition probabilities of the airports resulted in a discrete evolution graph that looked like a Markov chain. The author identifies different paths of airports due to a variety of factors involved in the historical variations for the 29 years. At last, the finalized airport groups showing similar evolutionary patterns are analyzed; airports with similar magnitude of network metrics (i.e., major, hub, regional) evolved through similar pathways.

To sum up, studies focusing on the components or functionality of the ATN are essential in researching the esoteric aspect of the ATN's complexity, which could not have been accomplished if the research had focused on only the network metrics.

OBSERVATION 3: The ATN evolution has been influenced by a multitude of internal and external events, while preserving the mainstream of the growth of the H&S structure.

Finally, Table 3 summarizes the papers reviewed in this section.

Table 3 – Summary of papers on ATN topology analysis.

Author	Network	Chronology	Metric (primitive)	Metric (advanced)	Focused property/event and approach
Albert, Barabasi	complex graph	evolution	clustering, degree	average path, spectral density	small-world, scale-free, preferential attachment
Gegov et al.	ATN (US)	evolution	clustering, degree, volume	shortest path	scale-free
Guimera, Sales-Pardo	ATN (Europe)	snapshot	nodes, links, clustering, degree	assortativity, congestability, modularity, synchronizability	small-world, community structure
Guimera et al.	ATN (world)	snapshot	degree	betweenness, modularity	Poisson distribution
Watts et al.	complex graph	snapshot	clustering, degree	shortest path	network decomposition
Erdos, Renyi	random graph	-	degree		P2P, H&S structure
Bonnefoy, Hansman	ATN (US)	snapshot	degree, operation, volume	flight weighted degree	scale-free, small-world, triads,
Bhadra, Hogan	ATN (US)	evolution	RPM, demand, operation, volume		network decomposition, small-world, scale-free
Wandelt, Xiaoqian	ATN (world)	evolution	degree, density, clustering, volume	betweenness, closeness	random walk
Neal	ATN (US)	snapshot	clustering, degree	modularity	scale-free, history, topology
Newman	complex graph	-	degree	shortest path, betweenness	scale-free, community structure
Billie, Kincaid	ATN (US)	evolution	operation		scale-free, multiple airlines
DeLaurentis et al.	ATN (US)	evolution	clustering, density, degree, strength	betweenness, eigenvector, shortest path	scale-free
Azzam et al.	ATN (Europe)	evolution	clustering, degree, links, volume	shortest path	scale-free, flight trajectory, delay propagation
Fleurquin et al.	ATN (US)	evolution	degree, operation		scale-free fitting, small-world
Wandelt et al.	ATN (world)	evolution	nodes, links, clustering, degree, volume	assortativity	P2P, H&S structure, routing, airline competition
Cook, Goodwin	ATN (US)	evolution	nodes, degree, demand, operation		scale-free, random attack on hubs
Cheung et al.	ATN (US)	evolution	nodes, links, clustering, degree	assortativity, betweenness, shortest path, resiliency	link reliability, network reconstruction
Sales-Pard et al.	ATN (world)	snapshot	nodes, links, degree, operation	affinity, modularity	hierarchical network organization, topology
Goetz, Sutton	ATN (US)	evolution	nodes, links, employment, operation, volume		geographic effects of deregulation
Reynolds-Feighan	ATN (US)	evolution	nodes, links, operation, volume	Herfindahl index, Gini index, Theil's entropy	deregulation, airlines' evolution, airport ranking
Jia et al.	ATN (US)	evolution	nodes, links, clustering, degree, volume	betweenness, closeness, efficiency, shortest path	scale-free, small-world, airport similarity & mixing pattern
Brueckner, Pai	ATN (US)	evolution	links, links, volume, operation, route distance		technological impact, P2P, H&S structure
Azzam	ATN (world)	evolution	clustering, degree	proximity measure	airport taxonomical categorization

2.2 ATN Logistics Optimization

2.2.1 ATN Logistics

2.2.1.1 General *p*-hub Problems

Logistics is another critical research field delving into networks. Fundamentally, logistics is about resource allocation under given circumstances. From the first formulated seminal research by O'Kelly, [64] the ATN logistics research has evolved to form a broad topic dubbed '*p*-hub (or *p*-median) problem' which seeks to find optimum network solutions by allocating p hubs – the number of hubs per travel between the origin and destination – while satisfying a series of constraints. The majority of hub location research has studied a restricted problem up to two hub stops at maximum [65,66,67] because most H&S networks in the world have two hub stops at best. Thus, this section will review mainly papers on *p*-hub problems and network optimization.

The first general form of a linear model was developed by Campbell [68], which evolved into a compact form in years later. [69] The following set of equations is the basic form of a typical *p*-hub problem. [64,70]

$$\text{minimize} \quad \sum_i \sum_j \sum_k \sum_l W_{ij} X_{ijkl} C_{ijkl}$$

$$\text{subject to} \quad \sum_k Y_k = p \tag{1}$$

$$0 \le Y_k \le 1 \text{ and integer } \forall k \tag{2}$$

$$0 \le X_{ijkl} \le 1 \quad \forall i, j, k, l \tag{3}$$

$$\sum_k \sum_l X_{ijkl} = 1 \quad \forall i, j, k, l \tag{4}$$

$$X_{ijkl} \le Y_k \quad \forall i, j, k, l \tag{5}$$

$$X_{ijkl} \le Y_l \quad \forall i, j, k, l \tag{6}$$

Here, X_{ijkl} is the demand fraction from origin i to destination j routing via $i \rightarrow k \rightarrow l \rightarrow j$ (Constraint 3). $Y_k = 1$ if location k is a hub and 0 otherwise (Constraint 2). W_{ij} is the actual amount of from i to j. C_{ijkl} is the sum of standard costs per unit demand from i to j (e.g., $i \rightarrow k$, $k \rightarrow l$, and $l \rightarrow j$). The above objective function represents the cost of all Origin-and-Destination (O-D) pairs in the network. Equation 1 enforces p hubs per route. Constraint 4 assures the completeness of X_{ijkl} with respect to W_{ij}. Constraint 5 and 6 ensure that demand travels via established hubs.

The introduced linear mathematical model by Campbell has become pervasive in the field of logistic optimization. This seminar research facilitated new inceptions of elaborated approaches to be developed afterward. For example, the introduction of a mixed-integer linear problem, [71] which heuristically explores the feasible space with smaller numbers and constraints than the traditional approaches. The elaboration of the p-hub problem was deployed in a variety of distinct directions. Hubs with limited capacity, [68] continuous p-hub location, [64, 72] multi-objective problems that sought to minimize/maximize other metrics (i.e., total traffic volume, total travel time, total resiliency/robustness etc.) in addition to minimizing the total network cost, [73,74,75] hub locations under uncertainty, to name a few. [76,77,78]

2.2.1.2 ATN Logistics & Optimization

Since the ATN was a good example of a real-world complex H&S network, many logistic scientists and airlines have also been interested in the logistic problems in air transport. To the author's best knowledge, the first logistic problem dedicated to the ATN was formulated by Brown. [79] This research focused on the specific function of the ATN: the

number of flights, ticket price, and enplanement routing via hub airports. Aykin introduced a framework to design an ATN with two network policies, the corresponding exact and heuristic solution procedures for the top 40 airports in 1989. [72]

Jaillet [80] formulates a generalized integer linear programming model that allows up to two hub airports per route. The O-D pairs of demand are generated by using a simple intercity-passenger-travel demand model. This research considers multiple aircraft types as well as their availability, as represented in the following equations of one-stop problem formulation:

$$\text{minimize} \quad \sum_i \sum_j \sum_k d_{ij} c_k y_{ij}^k$$

$$\text{subject to} \quad f_{ij} x_{ij} + \sum_h f_{hj} x_{hij} + \sum_h f_{ih} x_{ijh} \leq \sum_h b_k y_{ij}^k \tag{1}$$

$$x_{ij} + \sum_h x_{ihj} = 1 \quad \forall i, j, h \tag{2}$$

$$x_{ij}, x_{ihj} \geq 0 \quad \forall i, j, h \tag{3}$$

$$y_{ij}^k \in Z_+ \tag{4}$$

Here, the objective function is trying to minimize the total network cost due to transportation via corresponding chosen aircraft type. d_{ij} is the great-circle arc distance between origin i and destination j, y_{ij}^k is the number of aircraft type k operated from i to j, b_k is the capacity of aircraft type k, c_k is the cost per mile of aircraft type k which serves the segment i and j (y_{ij}^k). f_{ij} is the total passenger demand from i to j and x_{ij} is the fraction of the demand which directly flies to the destination j, whereas x_{ihj} is the fraction that takes one-stop flights by passing through hub h. Constraint 1 is the number of available

aircraft type k, which plays a role as the bandwidth of segment i-j: maximum allowable number of passengers who can be served via the segment i-j. Constraint 2 ensures that every demand is accommodated by the airline. Under some reasonable assumptions and rationalization, the authors succeed in formulating a set of general airline network logistic design problems involving the fundamental components of the ATN – demand, aircraft type, and airports – as well as transforming the ATN logistics problem as the struggle for the airline to deploy its network policy in the real world. The detailed models and methods for defining components to tackle the ATN logistics can be found in Song's research (1995).

Campbell [81] proposes a hub-arc location problem to reduce the flow costs by transforming the conventional p-hub problem and applied it into the ATN logistics in conjunction with a detailed introduction of a dedicated solution algorithm specialized for the hub-arc problem in the authors' companion paper. [82] Concerning multiple objectives is recently an active realm of research field involving multiple goals to fulfill. [83,84,85] Costa [74] formulates a bi-objective ATN p-hub problem to simultaneously minimize the total transportation cost and time to process the flow entering hub airports. Considering uncertainty [78,86,87] has also been extensively studied. Uncertainty was mostly imposed on demand and traffic costs. Marianov [88] introduces a modeling method to quantify the uncertainty of the congested hub airports in the ATN. The author models hub airports into an M/D/c queuing system, i.e., Poisson arrivals, deterministic service time, and c airlines where a probabilistic approach is harnessed to describe the instantaneous states of hub airports.

Furthermore, in freight aviation networks, the logistics problems are handled in a different paradigm compared to the ATN. Because there is no restriction on the minimum number of the hub airports in the ATN, the parcel delivery companies such as UPS and FedEx strive to minimize not only the cost but also the number of hub facilities in the ATN. Armacost [89] formulates and solves an ATN logistics problem for the specialized next-day overnight delivery in collaboration with UPS. The problem simultaneously tackles aircraft routes, fleet assignment, and routing altogether for minimal cost. O'Kelly [90] extracted the fuel burn information from the public data source about FedEx air freighter activities by setting and solving a logistics optimization problem with seven hubs, including the Memphis facility. The identified aircraft-wise fuel burn is then related to the quantifiable measures of the costs allocated to the hub facilities. Intriguingly, the author performs a reverse-engineering the individual values of the optimum decision variables from the optimal results existing in the real-world rather than solve the defined problem.

In conclusion, ATN logistics optimization has been an important research area since numerous complex network optimization problems involving different factors, and constraints can be efficiently solved by employing appropriate programming algorithms. The value of the logistics research is that it allows network scientists to tackle complex hub-location problems, naturally leading to the strong H&S structure in the ATN considered in many cases. Finally, Table 4 summarizes the papers reviewed in this section.

OBSERVATION 4: ATN logistics optimization researches enable complex H&S-structured ATN topology to be designed by numerically solving complex optimization problems under the given circumstances.

Table 4 – Summary of papers on ATN logistics & optimization.

Author	Network	Focus	Solution	Objective (min)	Comment
O'Kelly	ATN (US, freight)	problem formulation	p-hub	cost	1st p-hub problem
Campbell, O'Kelly	Logistics	paper review			
Garcia et al.	Logistics	algorithm development	p-hub	cost	new branch-and-cut algorithm
Campbell	Logistics	paper review			
Skorin-Kapov et al.	ATN (US)	problem formulation	p-hub	cost	
Campbell	Logistics	problem formulation	p-hub	cost	
Ernst, Krishnamoorthy	Logistics	algorithm development	p-hub	cost	exact & heuristic algorithm
Aykin	Logistics	problem formulation	p-hub & routing	cost	demand consideration
Brimbert, ReVelle	Logistics	problem formulation	plant location	cost + profit (max)	
Costa et al.	Logistics	problem formulation	p-hub	cost + process time	capacitated hub, single-hub
Shin et al.	Logistics	problem formulation	p-hub & routing	cost (hub, spoke, vehicle)	two-layered, genetic algorithm
Makui et al.	Logistics	problem formulation	robust p-hub	cost (hub set-up, transporting)	capacitated hub, uncertainty (demand, process time)
Sim et al.	Logistics	problem formulation + algorithm development	stochastic p-hub	cost	service-level constraint
Alumur et al.	Logistics	problem formulation	p-hub	cost (hub set-up, network)	uncertainty (set-up cost in hub, demand)
Brown	ATN (US, commercial)	problem formulation	aircraft cost	profit (max)	logistics cost modeling in ATN design
Jaillet et al.	ATN (US, commercial)	problem formulation	fleet assignment	cost	early work of fleet assignment on ATN design
Campbell et al.	ATN (US, commercial)	problem formulation	p-hub, number of segments	cost	up to 4 hubs per route
Koksalan, Soylu	Logistics	problem formulation + algorithm development	p-hub	cost (network, inter-hub)	implementation of genetic algorithm
Mahmutogullari, Kara	ATN (US, freight)	problem formulation + algorithm development	p-hub	cost	hub location with competition, duopoly network
Soylu, Katip	ATN (US, commercial)	problem formulation	p-hub	cost + number of 2-stop routes	trade-off between operation cost & passenger satisfaction
Yang	ATN (China, freight)	problem formulation	p-hub & routing	cost	stochastic approach with uncertainty (demand)
Contreras et al.	Logistics	problem formulation + algorithm development	p-hub	cost	stochastic approach with uncertainty (demand, cost)
Marianov, Serra	ATN (US, commercial)	problem formulation	p-hub & routing	cost (hub set-up, network)	consideration of airport capacity & number of runways
Armacost et al.	ATN (US, freight)	problem formulation	fleet assignment & routing	cost	airline p-hub problem on real-world problem
O'Kelly	ATN (US, freight)	problem formulation	demand extraction	cost	reverse engineering of demand from real-world logistics data

2.3 ATN Topology Design

This section reviews studies on designing the ATN, especially either using different methods from logistic optimization or explicitly focusing on the fundamental components of the ATN: aircraft, demand, airport. One could harness the retrieved knowledge from the

established ATN (e.g., the analyses in the previous section) to simulate the seemingly chaotic manifestation of the ATN.

2.3.1 Theory-Based Model

As for the general transportation network design, traditional approaches in network design focused on how to distribute demand throughout nodes: such approaches include entropy model, [91] gravity model, [92] and logit model. [93] These models were mainly interested in how the network would grow in response to the given socio-economic circumstances while failing to account for constraints such as finite resources and validating the emerged network.

Kotegawa [94] developed and compared three representative algorithms: the logistic regression model, the fitness function model, and the surrogate-based model. The logistic regression is a probability density function fitted to historical data to forecast the probability for a new route to be created. Fitness function models explicitly engage the fundamentals of preferential attachment, explaining that the more critical a node is, the higher the probability of creating connections increases. Lastly, an Artificial Neural Network (ANN) model is developed for 244 airports. Among the three approaches, the ANN model showed the best accuracy, but the authors emphasize the difficulty of extracting insights from the ANN model that is akin to a black box.

Takahashi et al. [95] study and discusses different forecasting methods of the ATN. The prediction methods include the ROC (Receiver Operating Characteristic) curve method, logistic regression and measures method, and four-step method, all of which are based on the network structure information obtained from statistical datasets. The results

emphasize the need to understand the aviation network growth. Among the utilized three methods, it is confirmed that the ROC method shows the best prediction accuracy, although modifying the prediction function's form and coefficients can further improve the accuracy of the logistic regression method.

2.3.2 Component-Based Model

There have been various modeling researches that strived to combine the components and network design problems. Hsu and Eie [96] interconnect the ATN design model with the dynamic fluctuation of jet fuel prices and explored the sensitivity of the network structure in response to the perturbation. In the proposed model, the 'reliability' of a route is probabilistically determined by the profitability of the route compared to the initial value. The case study between the international Taiwan ATN shows that the reliability of the route increases as the load factor increases. The results of the study provide an insightful way to improve ATN routing, especially under the uncertainty of jet fuel prices.

Hu and Paolo [97] propose a Genetic Algorithm (GA) based on complex networks theory for designing an optimized ATN with various objectives. The GA embraces the complex network concepts and techniques to model the ATN relevantly. The paper focuses on developing a highly efficient crossover operator. In order to set up the network optimization problem, this paper defines and uses specific network metrics such as robustness, degree, and shortest path. In an example experiment, the algorithm shows that the resultant European ATN topologies for different design objectives.

ATN topology design researches involve the components of the ATN. Firstly, aircraft performance specifications such as capacity, speed, and range can cause a massive

disruption in ATN as introduced in the early work by Jaillet's research above. [80] Later, ATN design problems integrated aircraft design & choice problems by setting the performance parameters as decision variables. [98,99,100]

Taylor [101] interconnects the aircraft sizing problem with logistics optimization by simultaneously considering the network routing, vehicle specifications, and operations for designing an ATN. In the objective function, the unit cost for each aircraft is explicitly calculated from its performance specifications. The integrated design problem formulation yields different network topologies under different objective functions: optimizing network with fixed aircraft types, optimizing aircraft types with fixed network topology, and concurrently optimizing both. This integrated ATN design approach yields the reduction of the total network cost, which can be interpreted as the enhancement of operational inefficiency.

In the case of the demand as the source of the ATN emergence, some considered its intrinsic uncertainty in ATN design [102] while others actively incorporated demand forecast models to ATN design. [103] Han [104] proposes an innovative method for designing an airline network by interconnecting the elasticity of demand with the airline network structure. The proposed optimization problem considers both the ticket prices and the profit of airlines that eventually determine the traffic volume from the network structure. For the case study, the author examines the China domestic network involving the top 15 major airports and demonstrated that hub locations tend to select airports with a larger volume.

Lastly, airports have been mostly considered as capacity constraint factors in ATN design problems. [105,106,107,108] There are three representative methods to alleviate the airport saturation issues to accommodate the drastically escalating air travel demand: constructing more runways, constructing new airports or divvying up traffics to airports in proximity, and increasing the number of night-time operations. The first and second can be easily incorporated in the ATN design study for making long-term and short-term policies.

2.3.3 Network Evolution Models

All literature papers reviewed so far strive to construct the target ATN topology from scratch, which is the *static* state. Many researchers, however, tackled designing ATN topologies from an evolutionary perspective. That said, there have been papers simulating the evolution of the ATN in response to model the decision-making mechanisms of stakeholders or policymakers.

Kotegawa [109] develops a network evolution model from the analysis of historical data. The author used three ML algorithms to identify evolution patterns of history. The used ML techniques are logistic regression, random forests, and support vector machines. The generated regression models of the evolution were used to model multi-layered demand, mobility, and capacity. The author finally compares and discusses the accuracy of each evolution based on different ML techniques.

Yang [110] develops a network design methodology that extensively harnesses the historical data to formulate the concept of evolution of the components in simulating the virtual network evolution. In the model, a two-dimensional convex path consisting of the number of airports and the fractional demand increment describes how the evolution of the

components is deployed under which the airline needs to construct the network topology by using a probabilistic distribution approach. To prove the proposed evolutionary concept, only the 53 top major airports are considered to incubate the growth of the primary H&S structure. To the author's best knowledge, this research is the first accomplishment to validate the simulated ATN to the real reference ATN.

OBSERVATION 5: ATN topology design researches focused on formulating the mechanisms of creating network segments by harnessing the known knowledge from analyzing the reference ATN.

Finally, Table 5 summarizes the papers reviewed in this section.

Table 5 – Summary of papers on ATN topology design.

Author	Network	Focus	Solution	Comment
O'Kelly	ATN (US, freight)	logistics (problem formulation)	p-hub	up to 2 hubs per route, interaction between hubs
Taylor, de Weck	ATN (US, commercial)	logistics (problem formulation)	aircraft design parameters	aircraft design + logistics
Potts, Oliver	Transportation	transportation (entropy model)		book on entropy-driven network construction papers
Bouchard, Pyers	Transportation	transportation (gravity model)		gravity-driven network construction approach
Teodorovic	Transportation	airline operations research		contain a paper on logistic regression for network construction
Kotegawa et al.	ATN (US, commercial)	comparison of design methods	minimum error to reference	fitness-based, logistic regression, artificial neural network
Takahashi et al.	ATN (US, commercial)	future forecasting of network	data-driven forecasting	operating characteristic curve method, logistic regression
Hsu, Eie	ATN (Asia, commercial)	logistics (problem formulation)	p-hub & fleet assignment	integration of fuel price to network design
Hu, Paolo	ATN (general)	logistics (problem formulation)	routing	integration of genetic algorithm
Teodorovic	ATN (Yugoslav, commercial)	logistics (problem formulation)	p-hub, aircraft sizing	minimum total flight time in a network
Mane, Crossley	ATN (US, commercial)	logistics (problem formulation)	p-hub, aircraft sizing	surrogate-based aircraft sizing + operation optimization
Crossley et al.	ATN (US, commercial)	network optimization	aircraft design parameters	aircraft design + fleet profit maximization
Burke et al.	ATN (US, commercial)	airline scheduling	optimum schedule	application of logistics in airline scheduling problem
Davendralingam, Crossley	ATN (US, commercial)	fleet assignment	aircraft design parameters	aircraft design + fleet assignment, test in real-world problem
Yang	ATN (China, commercial)	logistics (problem formulation)	p-hub & routing	stochastic model, uncertainty (demand)
Davendralingam, Crossley	ATN (US, commercial)	fleet assignment	aircraft design parameters	aircraft design + fleet assignment + demand forecast model
Han, Zhang	ATN (US, commercial)	parametric network design method	design parameters	up to 1 hub per route, city welfare, demand elasticity
Yang	ATN (US, commercial)	logistics (problem formulation)	p-hub & routing	up to 2 hubs per route, airport capacity
Davendralingam, Crossley	ATN (US, commercial)	fleet assignment	aircraft design parameters	robust aircraft design + fleet assignment + airport capacity
Wu, Zheng	ATN (general)	airport capacity constraint	p-hub	logistics with uncertainty (airport capacity)
Mohri et al.	ATN (Iran, commercial)	logistics (problem formulation)	p-hub	H&S design, airport capacity envelope, test on real-world problem

2.4 Research Statements

2.4.1 Research Gaps

Note that the research gaps will consider only the ATN logistics optimization and the ATN topology design in the literature review. It is because these two sections deal with papers about how to obtain an ATN, eventually. From now on, section 2.2 and section 2.3 will be

abbreviated as ‘ATN modeling’ research, while section 2.1 will be referred to as ‘ATN analysis.’

The first gap is that there are few specific implementations of the knowledge from analyzing the ATN or other types of networks in modeling ATNs. For example, as Neal [33] confirms, the ATN consists of multiple sub-networks having different rules. This discovery is considered in no studies. One discernible difference between ATN analysis and ATN design is that the realism of the target network; the former studies real-world ATNs, whereas the latter creates ATNs with little validation. This disjointed relation primarily originates from the difference of interests of the researchers.

In the architecture model’s perspective, however, this gap can be observed differently by addressing the following what-if questions:

- What if the design architecture can embrace the building blocks of the form and function identified by analyzing the ATN?
- What if the architecture model allows researchers to simulate a variety of different evolutions of the ATN by exploring the design parameters elaborated from absorbing what has been retrieved from studying the ATN?
- What if the architecture model can create numerous virtual ATNs where the over-arching form and function similarly manifest those of the real ATN with variations in minor characteristics such as the networks of different years?

Recalling the studies which analyzed the temporal evolution of the ATN, the underlying core characteristics such as scale-free, small-world properties and H&S structure for the last decades have always been dominant; variations were invoked by the gradual change

of various endogenous and exogenous circumstances. From this point of view, if the questions are relevantly answered, various virtual networks that will evolve from different configurations of modeling parameters of the model could have their model veracity, which will be analyzed, explored, and interpreted again. Then, the assets obtained from this analysis can be harnessed to elaborate and fine-tune the architecture of the model.

GAP 1: There have been few studies that integrated the ATN analysis with ATN design & modeling for the synergistic effect.

The second gap is the lack of effort to model evolution. In the papers of ATN modeling literature, every network is created, with no consideration of evolution since a logistics optimization is primarily a matter of how to allocate resources under given circumstances to either maximize or minimize the objectives. As for the ATN topology design research, most papers are focusing on integrating the components of the ATN into logistics problems to enhance the realism of the design architecture. As reviewed, most endeavors have been concentrated in aircraft fleet optimization or network optimization under additional constraints associated with the components (e.g., airport capacity).

GAP 2: Established ATN modeling researches have barely considered the evolution of the ATN both in the components and rules.

The third gap arises from the consideration of competition of airlines. In history of the ATN evolution, the structural transformation from P2P to H&S was accompanied by lots of Mergers and Acquisitions (M&A) between airlines. Their adaptive struggles for survival have driven the ATN evolution. Conceivably, the gigantic and complex ATN has

been constructed not by an omnipotent architect's will, but by the borderless blending of macro-and-micro-scale wrestling of airlines. However, this discussion is missing in the established ATN modeling literature papers; ATN logistics problems assume that there are no exogenous stakeholders. The problems assume an ultimate network constructor takes full control of all resources. Thus, in logistics problems, any airport can be a hub in order to find the optimum value of objective function (e.g., a network topology showing the optimum metrics of interest) as for ATN design problems, the propagation of network segments (the connection between nodes) is performed by a handful of universal mathematical principles because they are interested in figuring out a simple yet elaborate mathematical rule to efficiently design a network.

GAP 3: ATN design studies have not put critical attention to modeling the actual network construction policies of airlines.

The fourth gap is that the ATN modeling research does not validate and verify the created ATN against reference ATN. The literature studies are performed as if they were case studies; the ATN logistics and optimization research is to identify the optimum network topology, and the ATN design research focuses on developing and improving different approaches. Therefore, the following associated gaps are also identified:

- Not being interested in what happens inside the algorithm
- Using subjective values of design parameters
- Resultant network without validation

Note that these gaps originate from the different research goals of the established ATN design researches. For example, most logistics optimization research assumes a unique network stakeholder. However, considering the importance of the validation of the network for fulfilling the research objective, these gaps must be filled. Finally, the following insight is obtained:

GAP 4: Established ATN modeling research seldom performs verification and validation of the resultant network topology for realism.

To sum up, the identified research gaps will play a role as the foundation for formulating the Research Questions (RQ) and Research Hypotheses (RH) of this thesis.

2.4.2 Research Statements

Four research gaps were identified based on the five observations made from the literature review. Eventually, the following research gaps were made:

1. There have been few studies that integrated the ATN analysis with ATN modeling for the synergistic effect.
2. Established ATN modeling researches have barely considered the evolution of the ATN both in the components and rules.
3. ATN design studies have not put critical attention to modeling the actual network construction policies of airlines.
4. Established ATN modeling researches have seldom performed verification and validation of the resultant network topology.

2.4.2.1 Research Question 1

The first research question and research hypothesis are stated below:

RESEARCH QUESTION 1

What is a better way of representing the evolutionary components to develop an architecture model that achieves a necessary level of realism?

RESEARCH HYPOTHESIS 1

A vector space that integrates the multi-dimensional evolutions of crucial components – airport, demand, and aircraft – established by extensive historical datasets with proper techniques can represent the historical deployment of ATN evolution with augmented realism.

The first research question addresses what needs to be fulfilled to implement the evolutionary mechanisms of the ATN in the proposed architecture model with improved realism. It is formulated based on the first and second research gaps; appropriate architecture for virtually nurturing the evolutionary mechanisms is essential to improve the realism of the architecture model. As a result, the first research question also brings forth the following sub research questions:

- RQ1.1: What criteria should be made to identify the 'evolving' entity in the ATN?
- RQ1.2: How can the multi-dimensional evolution of components be orchestrated to establish a realistic evolutionary environment?

- RQ1.3: How can the data or information of the early ages of the ATN evolution be obtained? And if there is little to refer to, what alternative approaches should be made?

RQ1.1 requires this thesis to clarify the scope and granularity of what is considered to evolve in the model architecture. RQ1.2 states the need for considering the integration of multi-dimensional characteristics of the evolution. RQ1.3 should be addressed because there is, naturally, no comprehensive dataset for reference so that the evolution of the early stages – roughly before 1980 – must be regressed.

It is important to note that augmented realism should be fulfilled by two different directions of the evolution: temporal and spatial. The retrieved datasets should be able to form a temporal sequence, where the current element affects the next. Therefore, a reasonable approach to estimate the missing datasets of the early years of the ATN evolution should be made. Moreover, the datasets should be spatially comprehensive to represent the entire ATN. Therefore, the data integrity of public references should be thoroughly investigated to establish the most reliable information for answering the first research question. By answering the first research question, it is expected that this thesis proposes a better way to consider the 'entirety' of the ATN evolution with a significant amount of realism enhancement.

2.4.2.2 Research Question 2

The second research question and research hypothesis are stated below:

RESEARCH QUESTION 2

Can the complex rules of airlines' network construction policy be modeled while capturing the critical characteristics of the strong H&S structure?

RESEARCH HYPOTHESIS 2

A multi-tiered network evolution approach where an aggregated single airline deploys different policies to construct the primary H&S network and the secondary network can adequately model the rules of ATN evolution.

The second research gap identifies that no established ATN modeling literature explicitly tackled the adaptive behavioral mechanisms of airlines; instead, a sole omnipotent network construction architect has been considered. Moreover, the second observation provides an insight that decomposing the ATN into its sub-tiers is an effective strategy to understand the ATN in detail. Conceivably, it can be interpreted as that different mathematical representation of rules can be formulated in different sub-network tiers. The related sub research questions are as follows:

- RQ2.1: Can the network construction policy of airlines be analyzed and explicitly formulated?
- RQ2.2: Can the complex airlines' competition be transformed into another simplified abstraction with the analogy?
- RQ2.3: Can the network be decomposed into its sub-tiers based on the dominance of different rules to model the mechanism of evolution better?

RQ2.1 is a pre-requisite question to address the applicability of grouping airlines with similarity. A variety of metrics of interest will measure the similarity: the composition of viable aircraft types, number of operating cities, to name a few. RQ2.2 addresses the possibility for the airlines to be merged into one giant super-agent architect with the relevant analogy. RQ2.3 concerns the existence of multiple sub-network tiers in the entire ATN.

The second research hypothesis has a different property compared to the established literature; most established ATN modeling literature papers were case studies. This thesis strives to mathematically represent the rules of airlines' network construction policy, which is focused on evolving the H&S structure in the real world. However, as the second observation reveals, there can be multiple sub-network tiers depending on different criteria. As such, the second hypothesis explores the possibility of representing the rules of ATN evolution by a combination of multiple sub-rules, each of which relevantly models the core mechanism, not by a single global one that mostly focuses on modeling the H&S structure. Once the second hypothesis is proved, then it will be possible to develop and apply more mathematical rules of network construction policies with proper elaborations in future research such as sub-H&S structures observed in the ATN such as that of ANC airport as a hub and its spoke airports in Alaska state.

2.4.2.3 Research Question 3

The third research question and research hypothesis are stated below:

RESEARCH QUESTION 3

How can the architecture model be verified and validated against the reference real-world ATN to ensure the achievement of a necessary level of realism?

RESEARCH HYPOTHESIS 3

Employing a set of representative metrics and methods from different networks to which the ATN belongs can test the validity of the simulated ATN. Considered network types are general graph, complex network, transportation network, and air transportation network.

The fourth research gap leads to the formulation of the third research question, which raises an important issue that is directly related to the research objective. The established ATN modeling research papers focused on obtaining the final network topology. Naturally, the lack of modeling the real-world ATN was led to verification and validation of the created network against its reference. Thus, even though the first and second hypotheses are successfully proved, it is still doubtful whether the developed architecture model can genuinely represent the reference ATN with a necessary level of realism both in its form and function. A such, the third research question also addresses the following sub-questions:

- RQ3.1: What criteria should be defined to validate and verify the simulated ATN comprehensively?
- RQ3.2: How can the exploratory and interpretative capability of the architecture model be tested?

RQ3.1 addresses the importance of formulating the relevant criteria for verification and validation of the model. The established ATN analysis literature will profoundly inspire the third research hypothesis. RQ3.2 considers what is expected once the third research question is answered. The third research hypothesis is also connected to the first research hypothesis in that the model's veracity will be identified through testing various criteria in the architecture model. All design parameters, evolutionary deployment information, and the used datasets will be based on real-world references. Multiple configurations could be identified as the candidates for validating and verifying the model. Once the third hypothesis is proved, then it can be concluded that the developed architecture model successfully represents the evolutionary components and rules of the real-world ATN. Then, the architecture model with a sufficient level of realism will be able to perform a variety of different virtual network evolution scenarios without losing the essence of ATN evolution. To that end, this thesis will perform a case study to forecast the future of the ATN disrupted by supersonic transports.

2.5 Research Scope and Assumptions

The following tuples organize the scope and assumptions of this thesis:

- Architecture & scope
    - All U.S. territories are considered.
    - Daily representative domestic civil ATN is considered.
    - The evolution of the ATN is re-enacted, but the model mainly concerns the present ATN.

- Chronological progression is represented in a series of notional discrete waypoints, each of which corresponds to a particular year between 1917 and 2018.
- Airlines are the actual architect that constructs the ATN under given environments.
- Airlines consider the sentiency of passengers against total flight duration in its operation cost.

- Assumption
  - Component: airport, demand, aircraft
    - Airports are ordered historically. If multiple airports are introduced in the same year, they are sorted by the total volume's descending order (preferential attachment).
    - Symmetric demand: passengers eventually come back to their origins.
    - Passengers' hourly income (monetary value of time) is identical.
    - Aircraft has only economy class seats.
  - Rules: airlines' network construction policy
    - Dichotomous treatment of demand for network construction
      - Primary tier: H&S structure for direct / 1-stop travels in major cities
      - Secondary tier: multi-stop travels in regional demands that do not meet the service viability

- Airlines are abstracted as an aggregated super-agent in conjunction with an analogy.
- Focus on node and segment: no consideration of aircraft routing (one aircraft type per segment).

Average daily operations: no timetable, no seasonality, no price fluctuation, constant stop-over time due to the shortage of data of references.

CHAPTER 3. FORMULATION OF EVOLUTIONARY COMPONENTS

CHAPTER 3 and CHAPTER 4 introduce the details about the mathematical foundations for the proposed architecture model. For better organization, these are conveyed by two dedicated chapters. CHAPTER 3 and CHAPTER 4 are committed to answering the first and second research questions, respectively. This thesis benchmarks the developed architecture model against the pre-established initial groundwork by Yang. [110]

3.1 Overview

In order to fulfill the research objective, the comprehensive evolutionary components, modeling the mathematical rules of the network evolution, and its validation & verification should be performed. Once the research objective is fulfilled, some modeling abilities can be expected as follows:

- The top-level architecture of the model abstracts the continuing struggle of airlines under varying circumstances: established/inherited ATN, airports, demand, and aircraft.
- Each sub-architecture is modularized so that a variety of methods having different efficiency, fidelity, and characteristics can be harmoniously integrated.
- A variety of ATN structures are created by adjusting the primary inclination of airlines towards direct, indirect routes or in-between with single or integration of multiple design architectures.

The universal flow of evolution must be able to provide an array of circumstantial environments, under which the airlines adaptively deploy their network construction strategies. Finally, the outcomes from the struggles would let the network emerge and evolve. Finally, Figure 15 delineates the network of the building blocks to convey how the whole methodologies construct the proposed architecture model in the end. As noted, readers can confirm that the minimalism approach is also implemented in abstracting the real-world complex ATN. Note that the terms V&V and ENV in the figure are the acronyms of the phrase 'validation & verification' and 'evolution environment,' respectively.

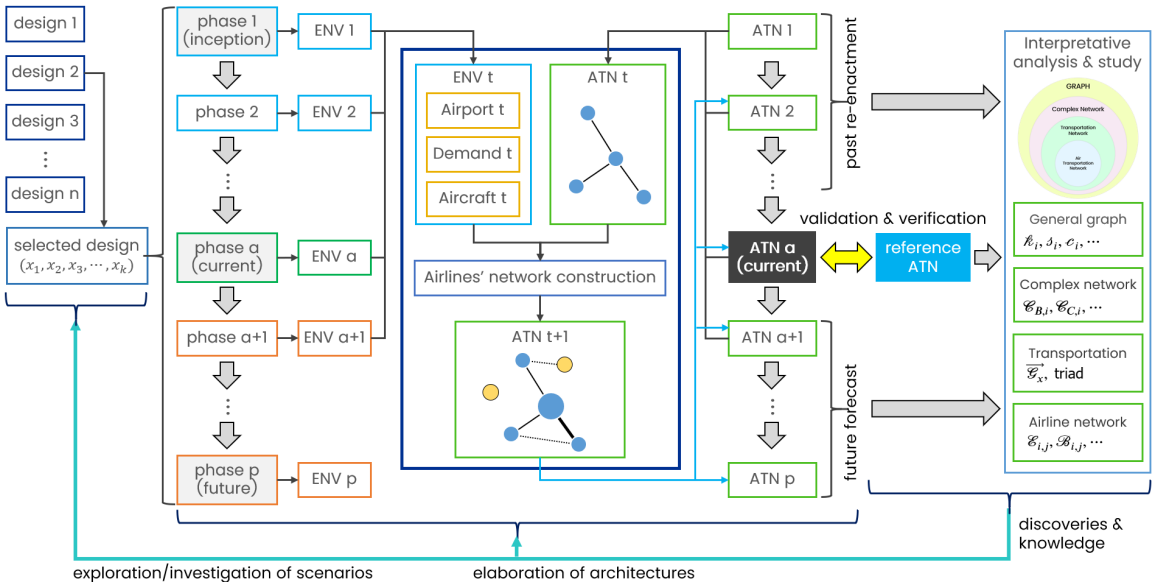


Figure 15 – Notional illustration of the developed architecture model.

As Figure 15 conveys, once the simulated ATN is validated and verified to successfully manifest the form and mimic the function of the reference ATN, then a variety of scenarios could be tested to enhance the level of understanding of its underlying

dynamics of the ATN. To recall the target level of capability for the evolution space to be equipped with, the first research hypothesis is revisited:

RESEARCH HYPOTHESIS 1

A vector space that integrates the multi-dimensional evolutions of crucial components – airport, demand, and aircraft – established by extensive historical datasets with proper techniques can represent the historical deployment of ATN evolution with augmented realism.

3.2 Foundation of Evolution Space & Evolution Path

The evolution of the air transportation network is the first item to tackle in this thesis. Tackling the intractability of the ATN puts one of its core tasks to transplanting the evolution, a physical phenomenon, into a cyber-physical space with short mathematical terms. To that end, the evolution is abstracted as a marching momentum that is established by the surrounding fundamental components of the ATN: airport, demand, and aircraft.

The modeling approach is decomposed into several sub-sections. First, a discussion will be performed on the history of ATN components to retrieve some insights about how evolution has been deployed. That said, the assumptions and scope of the design architecture will be outlined. Second, for each fundamental component, the details of research for grasping the mature dataset to describe what pathways it has been through so far aptly. Third, the independently identified dimensions of the evolution will be integrated to form the eventual evolutionary environment, which will have been capable of virtually germinating and raising the evolutionary momentum of the developed model, followed by the summary and discussions of this chapter.

3.2.1 Identifying Fundamental Components of ATN Evolution

3.2.1.1 Components from *Form*

A graph $G = (N, L)$ consists of nodes and segments. In ATN's perspective, nodes are airports, which relate to each other by abundant actual flights enabled by airplanes. Airports are responsible for the spatial expansion of the ATN, while aircraft connect airports. Considering the evolution, it is certain that airports have not inaugurated their civil aviation service at the same time and that the performance of aircraft has been advanced over history. If there had been no aircraft that can fly over the range of a segment, the direct segment could not have been made. From this perspective, knowing that all airports are not changing their geographical locations, the history of debuts of the airports into the civil aviation market, and the history of the technological advancement of various aircraft types are entirely correlated to each other.

To sum up, the ATN could have evolved in different pathways in response to either or both two components. Therefore, to take the form of the ATN to come to its existence, airports and aircraft must be considered. Figure 16 represents snapshots of the forms of the evolution which emerge under different evolutionary states of airports and aircraft.

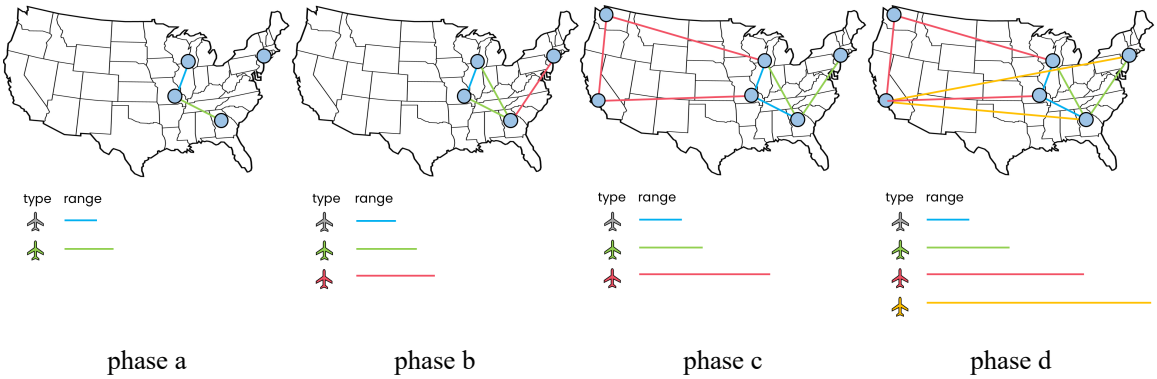


Figure 16 – Notional snapshots of ATNs under interlocked airports and aircraft technology evolutions.

Note that phases are arbitrary, and every representation is notional. At each figure, the horizontal lines under the 'range' text represent the range of each aircraft type. From the phase a to phase d, not only are the lines gradually lengthened but also new aircraft types are introduced, implying the aircraft technological evolution. Once the enablers become capable of directly connecting two airports in a far distance, the direct flight is started to appear.

3.2.1.2 Components from *Function*

The function is the behavioral dynamics of the form. Its effect is embodied in many ways and affects the form, which can be measured by various metrics such as volume, centrality, bandwidth. In a nutshell, the primary topic of the function of the ATN concerns what/who flows through the segments and what/who constructs the network. The former and the latter indicate passengers and airlines, respectively.

More specifically, passengers are abstracted into the term demand, which is the source and root of the ATN, and airlines are abstracted into the representative architect who allocates their resources to accommodate the passengers with minimum cost locally

and globally. As noted in the literature survey, numerous ATN modeling problems commonly have the premise that there is an amount of demand or commodities to travel from one place to another, and the active agent should find the optimum network topology to accommodate the need best.

The same analogy with this inference can even be found in what has been discussed in section 3.1. Figure 15 can be abstracted, in the most simplistic perspective, as a function mapping demand to ATN: ATN $= f(demand)$. In this subject, the actual performer is airlines, whose struggles create complex aspects of the *function* of the ATN. Therefore, demand and airlines are also the fundamental components of the ATN evolution. Figure 17 notionally illustrates the organization of this section.

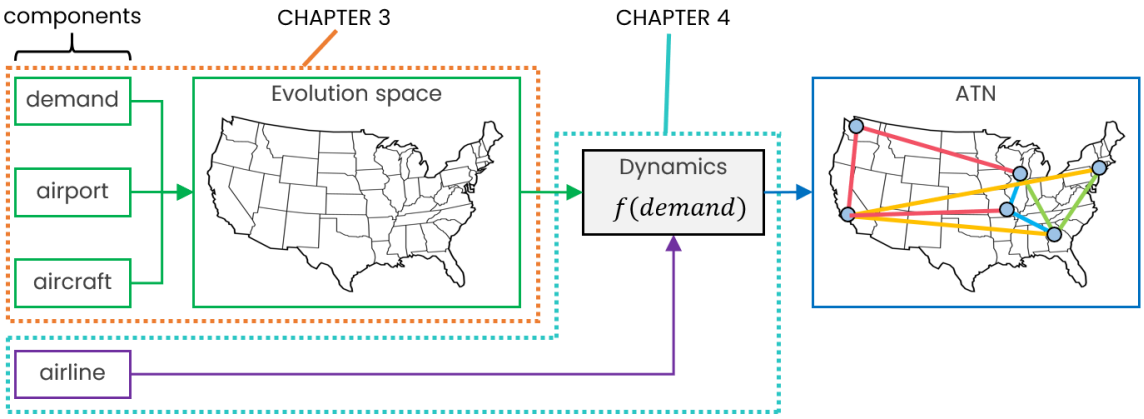


Figure 17 – Fundamental Components of ATN Evolution.

3.2.2 Evolution Space & Evolution Path

The identified components deploy their evolution in three fundamental dimensions. At any historical moment, each can be mapped to a particular 'point' describing the state, connecting them will result in a path. Finally, a virtual vector space is envisioned to mathematically represent an arbitrary 'path' connected by 'points' accounting for the three

fundamental components: a set of airports (A), a matrix of trip demand (T), and a set of aircraft types (Ψ), all of which depend on a discrete time step t as shown in the following equation:

$$\mathrm{ENV}(t) = \{A(t), T(t), \Psi(t)\}, \tag{3.1}$$

where

$$A(t) = \{\alpha_i(t) | i = 1, 2, \cdots, |A(t)|\} \tag{3.2}$$

$$T(t) = \left(\tau_{i,j}(t)\right) \in \mathbb{R}^{|A(t)| \times |A(t)|} \tag{3.3}$$

$$\Psi(t) = \{\psi_m(t) | m = 1, 2, \cdots, |\Psi(t)|\} \tag{3.4}$$

The evolution environment at time step t $\left(\mathrm{ENV}(t)\right)$ is defined by $|A(t)|$ airports and $|\Psi(t)|$ aircraft types which are available with a trip demand prescribed by $\tau_{i,j}(t)$, the number of passengers who want to travel from airports $\alpha_i(t)$ to $\alpha_j(t)$. As a result, the interconnection of all elements of $\mathrm{ENV}(t)$ would draw a curve, dubbed 'evolution path,' in the evolution space.

3.3 Study and Analysis of the Components

3.3.1 Airport (A): Node of ATN

3.3.1.1 Considered Airports

Several public sources are containing the inauguration year for civil aviation and geographical information (e.g., longitude and latitude). Thus, the first item to conduct is to build a list of considered airports. In order to do this, three independent datasets are engaged and cross-compared:

- DB1B: Airline origin and destination survey [111] – A 10% sample of airline tickets from reporting carriers collected by the Office of Airline Information (OAI) of the Bureau of Transportation Statistics (BTS).
- T-100D: Domestic Air Carrier Statistics (Form 41 Traffic) [7] – An air carrier statistics database that contains domestic airline segment data. The OAI of the BTS collects the data.
- PBACD: Passenger Boarding and All-Cargo Data for U.S. Airports [112] – A database provided by the Federal Aviation Administration (FAA) about passenger enplanement and cargo data extracted from the Air Carrier Activity Information System (ACAIS).

The DB1B and T-100D contain all origin-destination information in the U.S., whereas the PBACD contains only the airport-wise information so that by combining the active airports from each of these three public databases, a complete list of airports in consideration can be obtained. Table 6 summarizes the number of airports each year with at least one passenger per year.

Table 6 – History of the number of active airports in references.

Year	FAA	DB1B	T-100D
1990	-	-	297
1995	-	480	365
2000	545	462	340
2005	514	420	473
2010	498	402	462
2015	503	450	464
2016	506	435	472
2017	511	436	475
2018	-	438	478

The last row in Table 6 represents the number of distinct airports for each year, while the last column is corresponding to the number of distinct airports for the whole years of interest. As a result, this thesis considers 438 distinct airports in total as of 2018. (i.e., $A(f) = 438$). Conceivably, each airport has its debut year so that there will be a different number of airports in different evolution steps.

3.3.1.2 Fundamentals of Airport Data: Inauguration Year, Longitude, Latitude, Capacity

In this thesis, the developed architecture model is supposed to begin its simulation from the earliest inauguration year among the considered airports (i.e., min(year)). Generally speaking, once the moment is determined, the longitude and latitude must be given to place an airport at the right place and at the right year. Note that the capacity is handled in CHAPTER 4 as it is closely related to the mathematical rules of the ATN evolution. Eventually, the evolution of the airports can be described by a histogram based on the dataset associated with all considered airports. Finally, the sequence of debuts of airports

is created so that the cumulative number of airports can be mathematically represented as a function of time step t: $A(t)$.

Table 7 tabulates the characteristics of each reference source for comparison. [113,114,115,116,117,118,119,120] It is remarked that the AirNav, FAA Form 5010, OpenDataSoft, and Airport-Codes are the primary references as the government-related institutions provide them. Other data sources are used only as supplementary for filling the gaps of data from the primary reference sources. If there is an airport from which none of those references has the historical information, then its dedicated sources such as airport's homepage, local publications, encyclopedia are used, and corresponding references are attached, accordingly.

Table 7 – Comparison of reference sources for airport information.

Resource	AirNav	Prokerala	FAA Form 5010	Airport-Data	OpenData Soft	Airport-Codes
Data Type	Text	Text	Text, PDF	Text	CSV	Text
IATA code	O	O	O	O	O	O
Longitude	O	O	O	O	O	O
Latitude	O	O	O	O	O	O
Debut Year	△	△	×	×	O	×

In Table 7, the symbols 'O', '△', and '×' mean almost 100% of the information, not all-inclusive, and not available, respectively. As seen in the table, all data sources contain the complete coordinate information, whereas the debut year information is available in some of them.

3.3.2 Demand (***T***): *Root of ATN*

Since the demand is a priori, it can barely be measured or surveyed. Only what can be obtained is the enplanement. Therefore, it is reasonable to parse itineraries for identifying the true origin & destination, and it is also reasonable to consider the number of passengers throughout the itineraries as the demand. The only public comprehensive source datasets are DB1B and T-100D by the BTS. Figure 18 shows the historical variation of DB1B breakdown decomposed into non-stop, one-stop, two-stop, and more-than-two-stop routes for all available DB1B databases from 1993 to 2018.

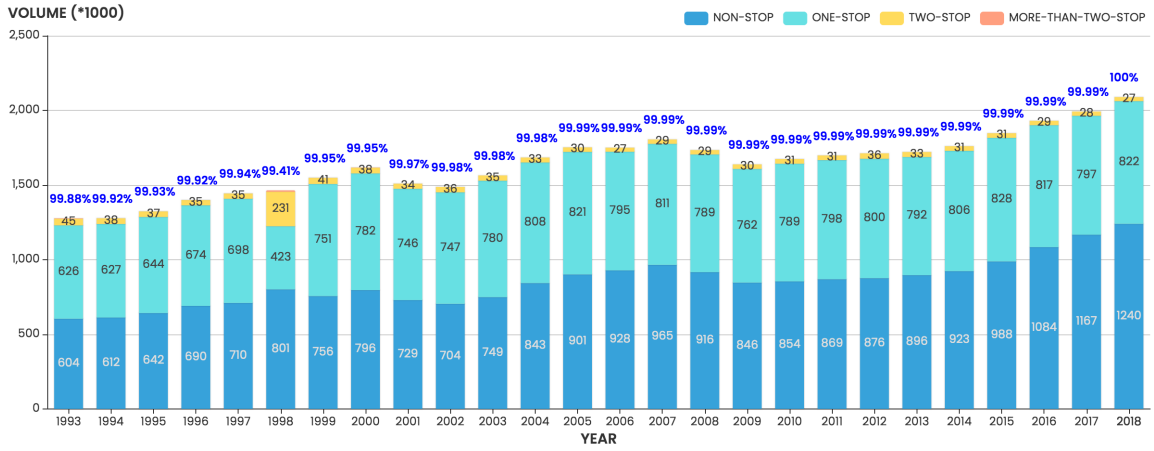


Figure 18 – Breakdown of DB1B daily volume from 1993 to 2018.

In Figure 18, the values inside stacked bars indicate corresponding kilo-people per day. The percentage value at the top of each stacked bar represents the total portion of the volume for non-stop, one-stop, and two-stop trips. For all years, they take over 99%, and the magnitude is even gradually increasing as time goes by. Especially for the year 2018, Figure 18 is construed as that approximately 2.16 million passengers take flights per day, and 99.99% of them travel via either non-stop, one-stop, or two-stop routes. Therefore, this thesis considers these three types of trips for modeling.

In the DB1B, identifying the full sequence of an itinerary can be performed by checking the 'trip break' information. Regardless of trip types (one way and round,) all final destinations of all routes are marked as 'trip break.' Moreover, according to the official documentation of DB1B by BTS, a flight is marked as 'trip break' if the flight angle difference between its adjacent ones (before and after) is over a certain amount of passengers are deemed to stay more than a certain amount of time at the airports with the trip break. This convention describes the final destination of a round trip marked as 'trip break' in most cases.

Therefore, by checking the presence of trip break information for all trip sequences in the DB1B, the full itinerary information can be possibly identified. Due to ambiguity from the absence of timestamp information, systematic sorting and filtering of the ticket data are performed focusing on direct, 1-stop and 2-stop flights of round trips. Moreover, separate one-way tickets are paired up with their counterparts to make round trips, assuming all the tourists eventually come back home. This process results in a subset of DB1B, dubbed SDB1B that contains only symmetric flight itineraries.

Two important detailed information can be retrieved from SDB1B with respect to each origin and destination pair from which the final O-D demand matrix $T(f)$. Since the matrix still accounts for approximately 10% of the total volume by its description, its magnitude is scaled up to match 100% volume of T-100D that is supposed to be a total enumeration. It is important to note that T is asymmetric by nature but possesses intrinsic characteristics granting reasonable extrapolation. [121] In contrast, volume (enplanement) is practically symmetric but extremely difficult to predict. As to the data structure, the O-D demand matrix consists of two sub-groups (i.e., primary and secondary) such that

$T(f) = T_{[P]}(f) + T_{[S]}(f)$, where $T_{[P]}(f) = \left(\tau_{i,j}(f)\right) \in \mathbb{R}^{|A_{[P]}(f)|\times|A_{[P]}(f)|}$ and $T_{[S]}(f)$ is the rest which is very sparse as the subscript '$[S]$' inadvertently indicates.

3.3.3 Aircraft (Ψ): Enabler of ATN

The history of the ATN is teeming with innovations on aircraft technology that have improved range, speed, fuel efficiency, and services. It requires no research statement to prove that different mixes of aircraft at different times impacted the course of the ATN evolution. It is, however, too complicated to consider all aircraft types with their performance characteristics, including cost information. In order to be considered, any aircraft type:

- must be able to obtain its cost information for both fixed and variable.
- must have a full set of significant performance specifications: speed, range, capacity.
- should have at least a certain amount of market share in terms of enplanement and operations.

The first two are indispensable, while the third one is preferable. Assume there are two different aircraft types: type 1 and type 2. Type 1 satisfies the first two requirements and has a tiny portion of market share, whereas type 2 meets either of the first two requirements and has a significant amount of market share. In this case, type 1 will be considered in this thesis.

There are three different datasets publicly available: T-100D, Schedule P-5.2, and supplementary materials from aircraft manufacturers and airlines. Schedule P-5.2 is

another aircraft-related database called air carrier financial data of Schedule P-5.2 from U.S. DOT From 41 by the BTS. [122] This database contains detailed quarterly operating expenses for large certificated U.S. air carriers. Table 8 summarizes the features of the considered aircraft data source.

Table 8 – Benchmark of Publicly Available Data Sources.

Feature	T-100D	P-5.2	Manufacturers & Airlines
Aircraft types	O	O	O
Utilization	calculable	O	X
Performance specification	O	X	O
Aircraft history	X	X	O
Cost (fixed & variable)	X	O	X

As categorized, the cost information can be only obtained from the Schedule P-5.2 while the history of aircraft can be acquired from the airlines' public data and other published materials. Therefore, the thesis is based on the aircraft types in Schedule P-5.2, which contains the cost information of them. The T-100D and other airlines' materials will be supplementary to the list of identified aircraft types from the Schedule P-5.2. This thesis does not calculate the exact fuel burn estimation along 4D flight trajectories to estimate the cost. Instead, collectively, the aircraft mission analysis will be extremely simplified to maximize the usability of the reference data aggregated from reporting carriers without delicate information such as instantaneous timestamps and weather conditions. As for the data sources from the aircraft manufacturers such as Boeing and Airbus, exploring their publicly released aircraft data allows researchers to retrieve not only the specific performance data but also the debut years in the history.

In Schedule P-5.2, variable costs are categorized as fuel and oil, maintenance, and crew. Fixed costs are categorized as depreciation, rentals, insurance, and others. Most cost categories are comprised of multiple items from Form 41 Schedule P-5.2, following the guide in the economic analysis of investment and regulatory decisions by the FAA. [123] One hurdle is that it does not contain the block time information. Therefore, ranking the aircraft types by their block time cannot be estimated. Due to this issue, this thesis uses Schedule P-5.2 from 2002 to 2017, from which the hourly cost can be estimated. It should be remarked that some aircraft types come with a lack of information so that if the cost data of an aircraft type is missing equal to or more than ten times, it is not accounted for the consideration of aircraft types.

In this thesis, the fleets of the four major airlines in the ATN are considered to explore: Delta Airlines (DL), American Airlines (AA), Southwest Airlines (WN), and United Airlines (UA). According to the T-100D 2018, these airlines take about 87% of the total network volume. Therefore, it is reasonable to consider only these top four majors. [124,125,126,127,128,129,130,131,132,133,134,135] One remark is made; the aircraft types introduced in the early 20th century have no financial data. In terms of evolution, however, those such as the Boeing 40A and Boeing 40B should be considered to re-enact the ATN evolution in the developed architecture model to grasp and augment realism. As such, these aircraft types are not to be excluded as those aircraft types in the Schedule P-5.2 with lack of data are omitted out of consideration.

3.3.3.1 Grouping of Aircraft Types Using K-means Clustering Optimization

It is challenging to retrieve the block time or utilization of all 56 aircraft types. Thus, this thesis abstracts those 56 aircraft types into a smaller number of representative groups based on the similarity in performance: capacity, range, and speed. To that end, they are populated into a 3D space, and K-means clustering is used to identify the appropriate number of groups. [136] K-means clustering is a popular unsupervised machine learning algorithm for dimensionality reduction in data science.

To identify the most desirable number of clusters, a popular statistics and data analysis tool JMP®[137] is used to solve an optimization problem whose goal is to minimize the average of standard deviations of all clusters such that $\left(\sum_{i=1}^{K}\sum_{j=1}^{3}\sigma_{i,j}\right)/K$, where $5 \leq K \leq 10$ is the decision variable, and j is the index of the criteria: capacity, range, and speed. As a result, optimum $K = 10$ is found. Table 9 summarizes the result. Note that if $K > 10$, then at least one cluster consists of one element, so the design range of K is set between 5 and 10. Also, Figure 19 and Figure 20 visualizes the optimum 10 clusters into a 2D and 3D scatter plots.

Table 9 – K-means Clustering Result.

K	Absolute			Normalized $(\sigma_{i,j})$			$\left(\sum_{i=1}^{K}\sum_{j=1}^{3}\sigma_{i,j}\right)/K$
	Capacity	Range	Speed	Capacity	Range	Speed	
5	23.28	357.52	43.74	1	1	1	1
6	17.54	296.23	33.96	0.430	0.560	0.648	0.546
7	17.74	252.98	25.86	0.450	0.250	0.357	0.352
8	14.99	252.27	23.99	0.177	0.245	0.289	0.237
9	14.48	218.10	22.65	0.127	0	0.241	0.123
10	13.21	231.34	15.95	0	0.095	0	0.032

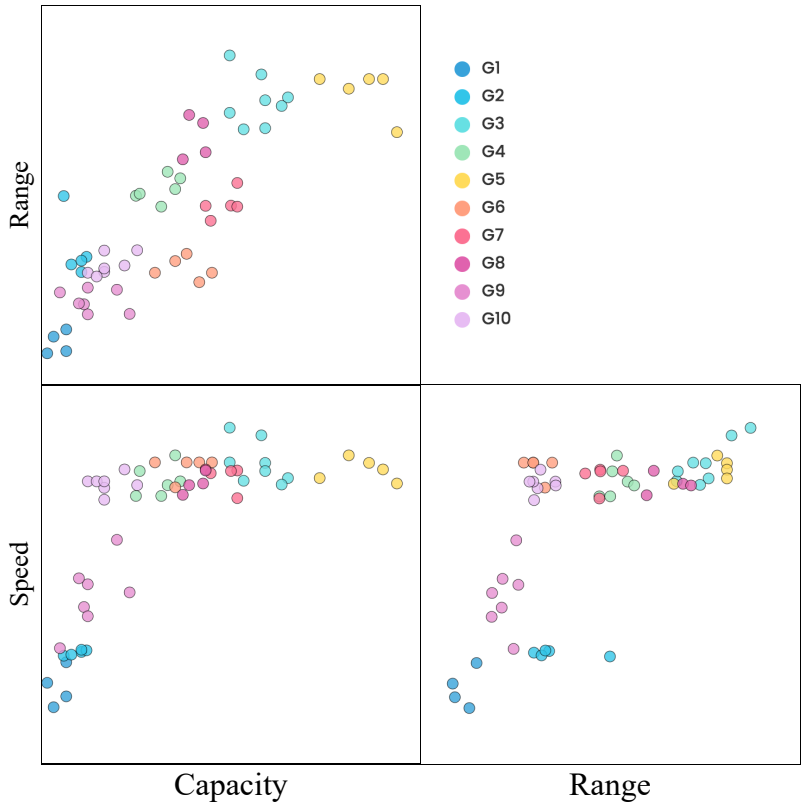


Figure 19 – Scatter plot matrix of the optimal ten aircraft groups.

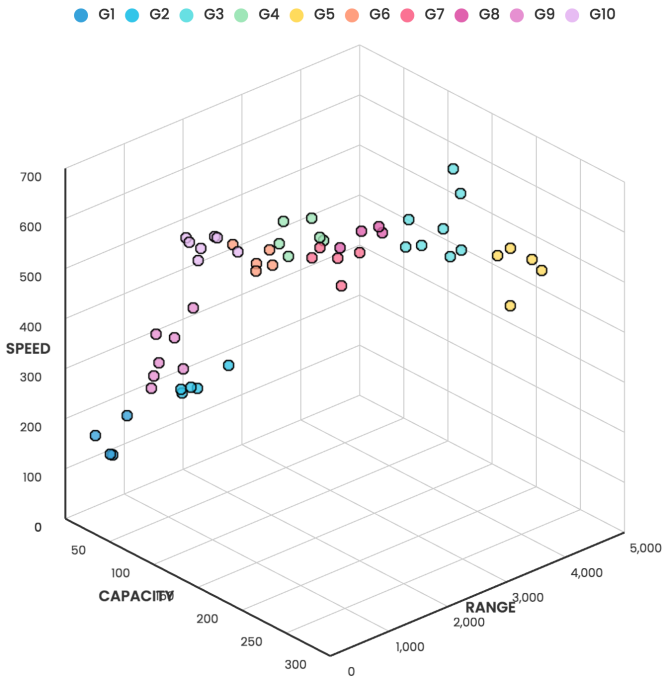


Figure 20 – 3D scatter plot of 56 aircraft types.

In Figure 19, the aircraft groups are well-identified typically in one with capacity as the abscissa and range as the ordinate. Through this abstraction, the number of aircraft types is reduced from 56 to 10. Each group plays a role as the representative aircraft type for all associated specifications. All these three attributes are the basis for airlines to perform fleet assignments. Eventually, the complete information of all considered aircraft types with their assigned groups is tabulated in Table 10.

Table 10 – Complete information of all considered aircraft types.

Group	Aircraft type	Debut year	Capacity (pax)	Range (mi)	Speed (mph)	Block time (ratio)	Cost ($/hr)
1	Boeing 40A	1927	10	650	105	0.25	1,700
	Boeing 80A	1928	20	460	125	0.25	1,600
	Boeing 247	1933	20	745	188	0.25	1,900
	Bell B-206A	1962	5	431	150	0.25	1,536
2	Douglas DC-3A	1936	32	1,500	207	0.35	2,700
	Lockheed Model 18 Lodestar	1940	18	2,500	200	0.13	2,732
	Vickers VC.1 Viking	1946	36	1,700	210	0.15	2,310
	Douglas DC-5	1949	24	1,600	202	0.25	1,960
	Saab 90 Scandia	1950	32	1,650	211	0.12	2,085
3	Boeing 707-120	1959	174	4,100	607	0.05	4,257
	Douglas DC-8	1959	177	3,760	556	0.05	4,130
	Boeing 720	1960	149	4,350	621	0.05	4,687
	Convair 990	1962	149	3,595	557	0.05	4,294
	Airbus Industrie A320-100/200	1988	195	3,798	528	0.3638	4,225
	Airbus Industrie A321	1994	190	3,685	516	0.0423	4,458
	Boeing 737-800	1998	160	3,378	523	0.3666	4,272
	Boeing 737-900	2001	177	3,393	541	0.0273	3,960
4	Boeing 727-100	1963	106	2,591	570	0.0068	24,539
	Boeing 737-100/200	1968	95	2,361	495	0.0579	4,416
	Boeing 737-500	1990	110	2,733	522	0.2452	3,986
	Emb-170/EMb-175	2004	75	2,504	495	0.1880	1,600
	Embraer ERJ-175	2004	78	2,531	541	0.3054	1,566
	Embraer 190	2004	100	2,819	515	0.1967	3,837
5	McDonnell Douglas DC-8-71	1967	259	4,039	556	0.0476	13,454
	Airbus Industrie A300B/C/F-100/200	1974	281	3,340	518	0.0178	7,525
	McDonnell Douglas DC-10-30CF	1980	270	4,039	544	0.0185	9,667
	Airbus Industrie A310-200C/F	1983	220	4,039	528	0.2231	14,610
	Boeing 757-300	2000	243	3,912	570	0.6930	5,572

6	McDonnell Douglas DC-9-15F	1967	90	1,491	557	0.0062	5,594
	McDonnell Douglas DC-9-30	1967	115	1,740	557	0.1870	4,452
	McDonnell Douglas DC-9-40	1968	125	1,367	557	0.0346	4,986
	McDonnell Douglas DC-9-50	1975	135	1,491	557	0.0979	4,734
	Boeing 717-200	2000	106	1,645	511	0.6742	4,435
7	Boeing 727-200/231A	1968	134	2,175	537	0.0328	9,407
	McDonnell Douglas DC9 Super 80/MD81/82/83/88	1980	155	2,672	542	0.5000	4,739
	Boeing 737-300	1984	130	2,371	544	0.3393	4,113
	Boeing 737-400	1987	150	2,374	541	0.0864	4,642
	McDonnell Douglas MD-90	1995	155	2,361	491	0.0414	5,318
8	Boeing 737-200C	1968	112	2,983	497	0.0053	5,341
	McDonnell Douglas DC9 Super 87	1987	130	3,076	542	0.0024	3,555
	Boeing 737-700/700LR/Max 7	1997	128	3,461	518	0.9817	3,526
	Airbus Industrie A-318	2003	117	3,567	515	0.0106	3,292
9	De Havilland DHC8-400 Dash-8	1983	60	1,268	414	0.2439	2,098
	Cessna 208 Caravan	1984	15	1,232	214	0.1892	1,155
	Saab-Fairchild 340/B	1984	34	1,076	290	0.1930	1,483
	De Havilland DHC8-100 Dash-8	1984	37	944	273	0.0350	1,731
	De Havilland DHC8-200Q Dash-8	1984	37	1,295	332	0.0477	2,056
	Embraer EMB-120 Brasilia	1985	30	1,087	343	0.1931	1,430
	Aerospatiale/Aeritalia ATR-72	1989	70	949	317	0.0982	2,857
10	Canadair RJ-100/RJ-100ER	1992	50	1,502	488	0.0207	2,082
	Canadair RJ-200ER /RJ-440	1992	50	1,548	510	0.3694	1,940
	Embraer-135	1995	37	1,491	522	0.0245	1,839
	Embraer-140	1995	44	1,442	522	0.0382	1,871
	Embraer-145	1995	50	1,783	522	0.2775	1,452
	Canadair RJ-700	2001	66	1,586	544	0.1746	1,884
	Canadair CRJ 900	2001	76	1,787	515	0.095	1,752

3.3.3.2 Estimating Representative Specifications of Aircraft Groups

Because the abstracted ten aircraft groups have different values of a performance specification, the fixed cost and variable cost must be modified accordingly. To that end, hourly cost ($) and block time are needed. The cost must be averaged based on the block time of each aircraft. The principle is simple; the more an aircraft type is utilized, the more reliable and important the cost becomes. Schedule P-5.2 also gives the block time.

Therefore, the representative specification of aircraft types denoted as X is calculated by merely identifying the centroid as follows:

$$X(t) = \sum_{i=1}^{N(t)} \left(X_i(t) \times Blt_i(t) \right) \Big/ \sum_{i=1}^{N(t)} Blt_i(t), \tag{3.5}$$

where i is the aircraft type index in an arbitrary aircraft type group at time t, Blt is the block time obtained from the Schedule P-5.2, N is the number of existing aircraft types in the group at time step t, and X_i is an arbitrary specification of the i-th aircraft type in the group, such as capacity, range, speed, and cost.

3.4 Evolution Space and Evolution Path

This section integrates all the aforementioned fundamental components of the ATN into a multi-dimensional evolution space that can incubate, nurture, and mature the ATN.

3.4.1 Evolution of Airport

There are 438 airports in civil aviation services according to the reference data currently. [7,8,111] As the airports are the nodes of the ATN, the evolution path of airports per year can be established. By mapping the years from 1917 to 2018 to the time variable t, the evolution of airports can be represented as a path by using the number of airports. Figure 21 illustrates the yearly increment of airports as well as the cumulative number of airports.

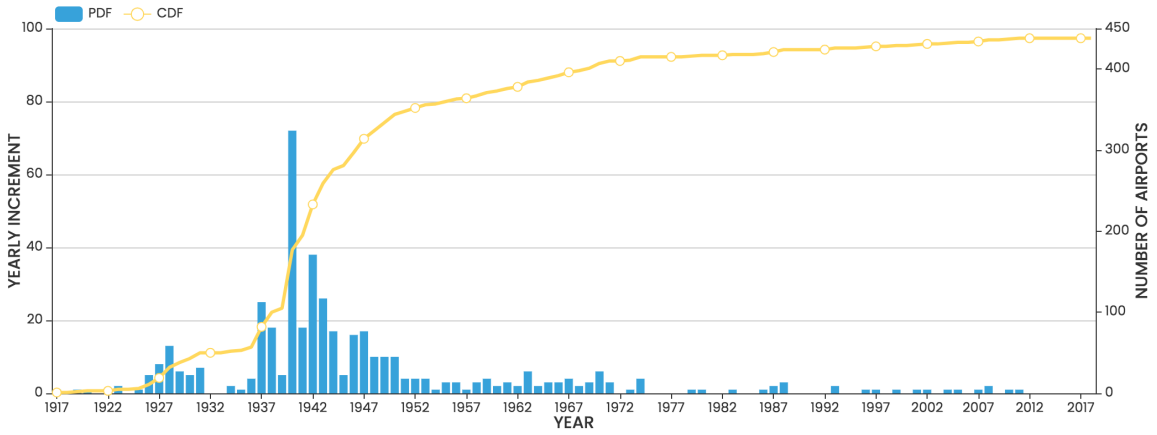


Figure 21 – Evolution of airports (PDF and CDF).

3.4.2 Evolution of Demand & Enplanement

3.4.2.1 Insights from Historical Observation

The earliest years of the available SDB1B and the T-100D are 1993 and 1990, respectively. As assumed, all components monotonically expand so that, arguably, there are no reasonable ways for reproducing the complete O-D demand and enplanement matrices for the years earlier than 1993. As such, this thesis harnesses the recent trends of the two databases to regress and model the O-D demand and enplanement matrices, albeit the methods being simplistic.

Since the modeling year is 2018 for verification and validation, there are two boundary conditions for the demand to satisfy: the beginning year and 2018. In the developed architecture model, the beginning year of simulation is 1919, as it is the first year when there are at least two established airports: DAL (in 1917, Dallas, TX) and TUS (in 1919, Tucson, AZ). Then, the trend of the ratio of the sum of demand to the sum of volume for the last 26 years (1993 ~ 2018) is analyzed to apply to the earlier years for

reasonably estimating the demand in the early years. To sum up, Figure 22 shows the historical variations of demand (SDB1B) and enplanement (T-100D).

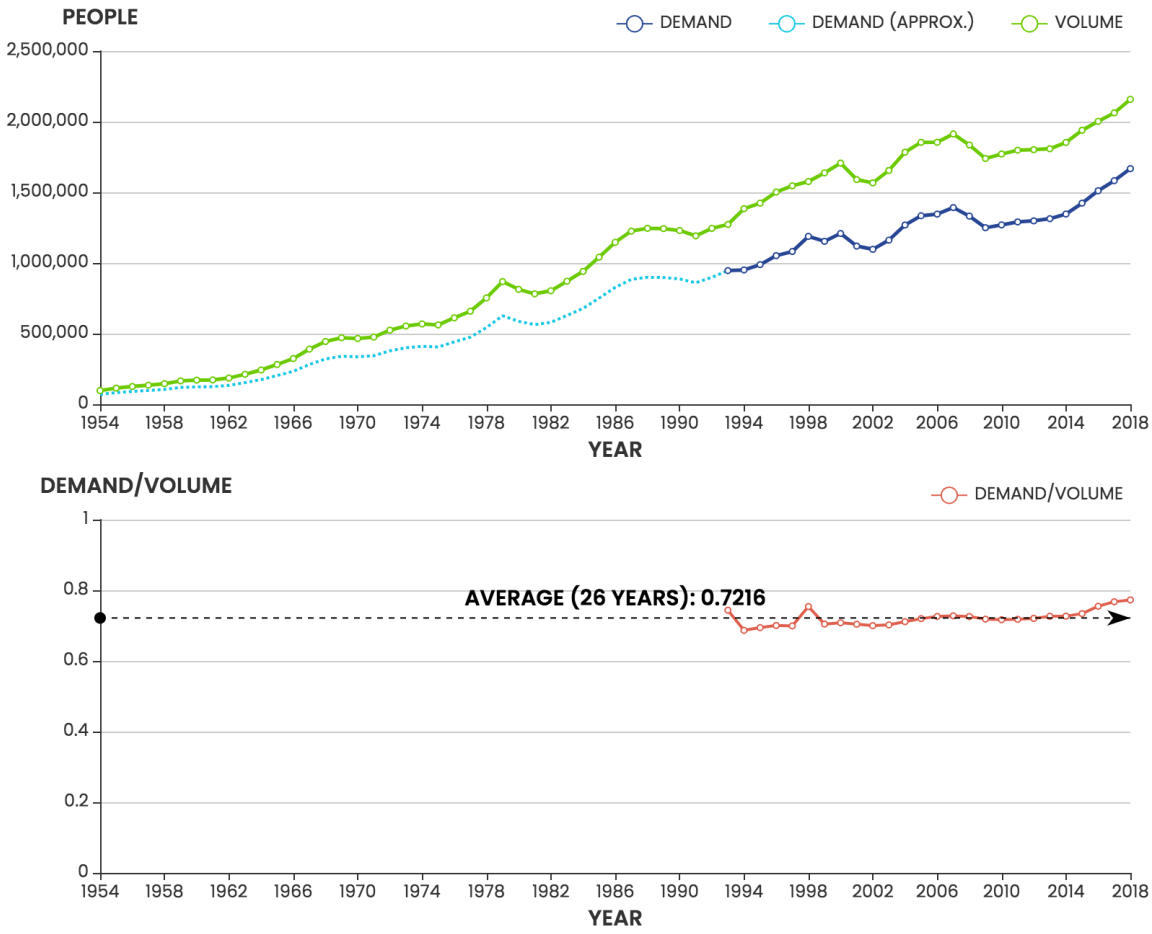


Figure 22 – Historical trend of demand and enplanement (1954 ~ 2018).

The dotted light-blue line in the upper chart represents an approximated value of demand extrapolated from 1992 back to 1954. In order to articulate the dotted line, the ratios of demand to enplanement between 1993 to 2018 were calculated. Its trend is represented in the lower chart showing its average 0.7216 throughout the 26 years (1993 ~ 2018) as well as its standard deviation of 0.0222. As the slight amount of standard deviation says, the light-blue approximated dotted line of demand is the result of what is multiplied by the average ratio 0.7216, which is the most straightforward approach. By using this ratio,

the demand history back to 1917 is approximated. Table 11 summarizes the approximated values of the sums of demand and enplanement, respectively, for several pivotal years before 1954.

Table 11 – Approximated value of demand and enplanement sums.

Year	T_{Σ}	$\mathscr{E}_{\Sigma}$
1917	72	100
1920	720	1,000
1930	14,400	20,000
1940	43,300	60,000
1950	81,200	112,000

3.4.2.2 O-D demand matrix

Once the evolution path of the demand is obtained, individual O-D demand should be defined. The evolution of demand is expressed in the following equation from the previous research by Lewe & Yang[121]:

$$\Delta\tau_{i,j}(t) = \tau_{i,j}(f) \times \frac{t}{N} - \sum_{k=1}^{t-1} \tau_{i,j}(k), \tag{3.6}$$

where $\Delta\tau_{i,j}$ is the change of demand, f is the final time step, and N is the number of time steps. In the previous research, [121] a limited approach was applied; only the positive increment of demand was considered so that a monotonic increment of demand evolution was assumed due to the algorithmic functionality. However, the real statistical data shows a complex mixture of increase and decrease throughout history. Therefore, this thesis expands the traditional approach in order to augment the realism for modeling the demand

evolution. Thus, the evolution of demand exactly follows the historical evolution path of demand in this thesis, as shown in Figure 23.

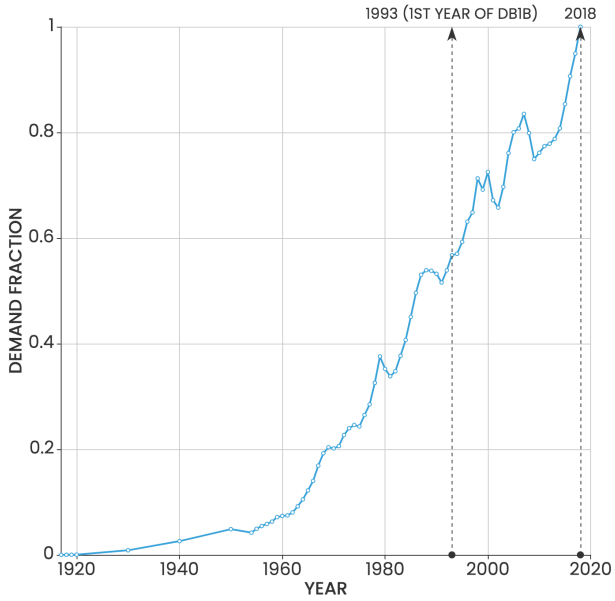


Figure 23 – Historical change of fractional demand (0 ~ 1).

3.4.3 Interlocked Evolution of Airport and Demand

The airport $A(t)$ and demand $T(t)$ are interlocked with each other. Recalling the definition of demand:

$$T(t) = \left(\tau_{i,j}(t)\right) \in \mathbb{R}^{|A(t)|\times|A(t)|} \tag{3.3}$$

As the equation prescribes, demand is only considered on the existing airports at time step t: $A(t)$. It is natural since $A(t)$ is of the spatial expansion of the ATN, whereas $T(t)$ is of the chronological progression of the ATN.

In summary, Figure 24 shows the evolution path we used in this study with graphic illustrations of the evolution of $A(t)$ and $T(t)$. In the next section, we will discuss the adaptive dynamics that is a collection of actual algorithms of the proposed model. Each bar represents $\tau_{i,j}(t)$ in the vertical axis while the two horizontal axes represent the origin (α_i) and the destination (α_j), respectively.

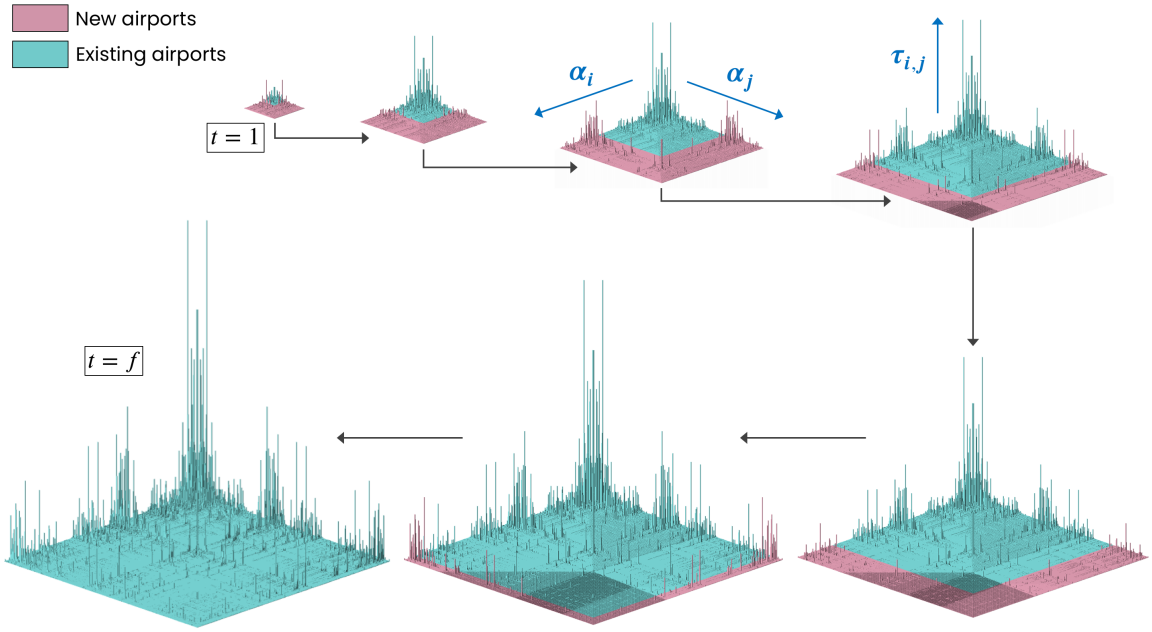


Figure 24 – Topological illustration of the evolution of $A(t)$ and $T(t)$.

3.4.4 Evolution of Aircraft

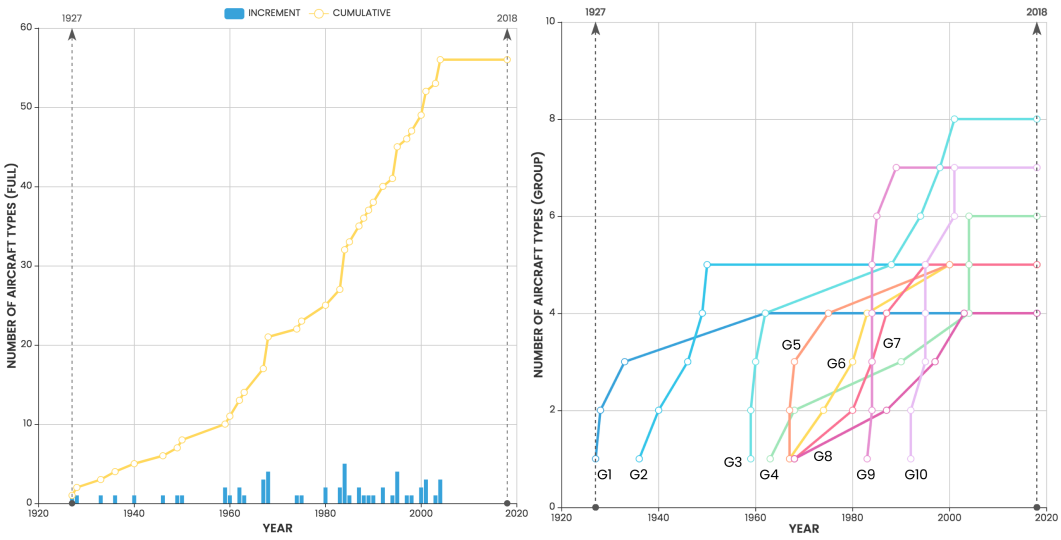
Each aircraft type group information contains an array of representative specifications: debut year in the market, speed, range, capacity, block time, hourly cost. All details of these properties are explained in section 3.3.3 so that the complete evolution of aircraft types can be formulated, in the long run.

Different aircraft types take their advent in a different moment in the timeline so that an advanced approach needs to be formulated to dynamically update the corresponding

specifications of all aircraft groups based on existing aircraft types within them. The basic principle consists of two considerations. First, the same approach accounting for the significance of contribution each aircraft type in cost is expanded to estimating performance specifications: capacity, range, and speed. Second, at every evolutionary moment, the relative contribution is adjusted so that the specifications of aircraft groups are dynamically updated. Finally, the complete evolutionary information of aircraft type groups is provided in Table 12 while Figure 25 and Figure 26 show the final evolution of the set of aircraft types denoted as $\Psi(t) = \{\psi_i(t) | i = 1, 2, \cdots, |\Psi(t)|\}$, containing the number of viable aircraft type groups.

Table 12 – Evolution of aircraft type groups.

Group	Year	Capacity (pax)	Range (mi)	Speed (mph)	Cost ($/hr)
G1	1927	10	650	105	1,700
	1928~1932	15	555	115	1,650
	1933~1961	17	618	139	1,733
	1962~present	14	571	142	1,684
G2	1936~1939	32	1,500	207	2,700
	1940~1945	28	1,771	205	2,709
	1946~1948	30	1,754	206	2,614
	1949	28	1,710	205	2,428
	1950~present	29	1,703	206	2,387
G3	1959	176	3,930	582	4,194
	1960~1961	167	4,070	595	4,358
	1962~1987	162	3,951	585	4,342
	1988~1993	183	3,852	548	4,266
	1994~1997	184	3,841	546	4,280
	1998~2010	175	3,666	537	4,277
	2011~present	175	3,659	538	4,268
G4	1963~1967	106	2,591	570	24,539
	1968~1989	96	2,385	503	6,531
	1990~2003	107	2,660	518	4,518
	2004~present	91	2,623	520	2,934
G5	1967~1973	259	4,039	556	13,454
	1974~1979	265	3,849	546	11,841
	1980~1982	266	3,891	545	11,362
	1983~1999	233	3,998	533	13,721
	2000~present	240	3,938	559	8,074
G6	1967	114	1,732	557	4,488
	1968~1974	116	1,676	557	4,564
	1975~1999	122	1,621	557	4,615
	2000~present	111	1,637	526	4,494
G7	1968~1979	134	2,175	537	9,407
	1980~1983	154	2,641	542	5,026
	1984~1986	144	2,536	543	4,671
	1987~1994	145	2,521	543	4,668
	1995~present	145	2,515	541	4,695
G8	1968~1986	112	2,983	497	5,341
	1987~1996	118	3,012	512	4,777
	1997~2002	128	3,458	518	3,536
	2003~present	128	3,459	518	3,534
G9	1983	60	1,268	414	2,098
	1984	38	1,192	315	1,658
	1985~1988	36	1,169	321	1,609
	1989~present	40	1,148	320	1,731
G10	1992~1994	50	1,545	508	1,948
	1995~2000	49	1,629	515	1,752
	2001~present	55	1,636	520	1,775



Aircraft types (full) Aircraft types (group)

Figure 25 – Evolution of aircraft types $(\Psi(t))$.

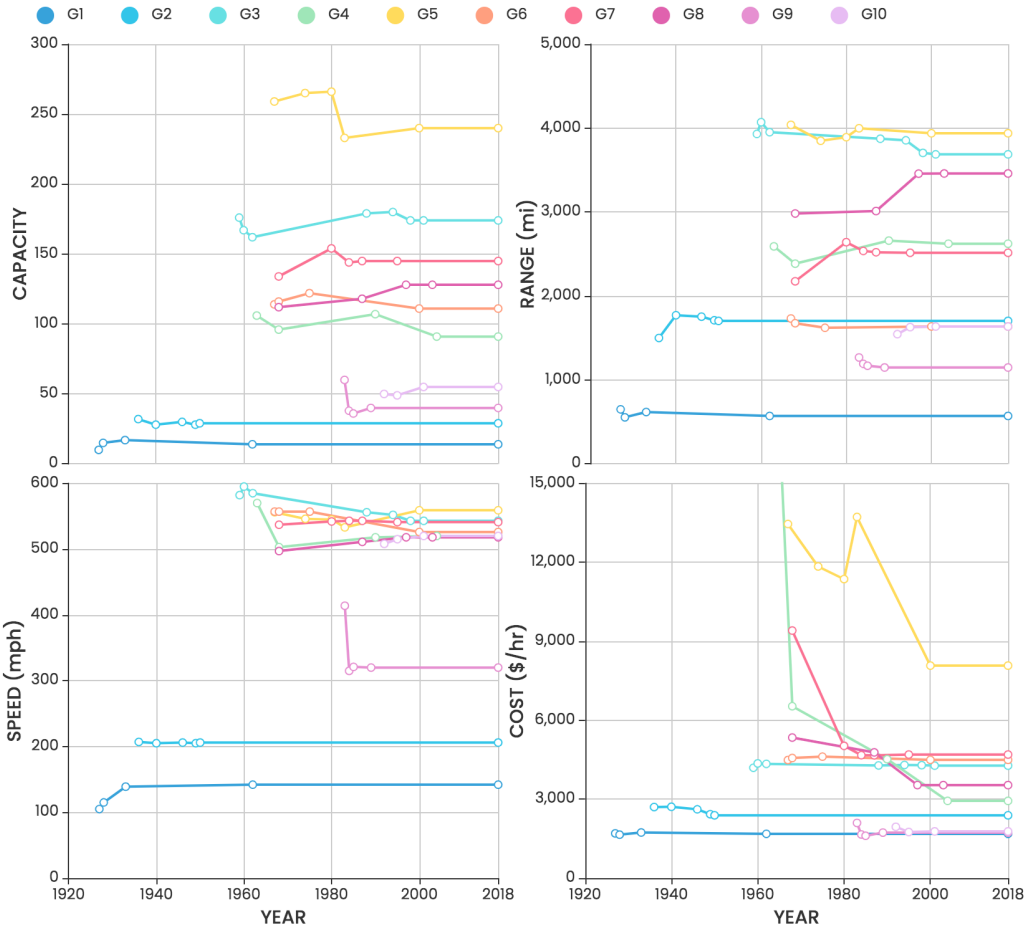


Figure 26 – Evolution of representative specifications of all aircraft type groups.

3.4.5 Evolution Space and Evolution Path

The proposed multi-dimensional evolution space is recalled; it consists of a set of airports (A), a matrix of trip demand (T), and a set of aircraft types (Ψ), all of which depend on a discrete time step t as shown in the following revisited equation:

$$\mathrm{ENV}(t) = \{A(t), T(t), \Psi(t)\}, \tag{3.1}$$

where

$$A(t) = \{\alpha_i(t) | i = 1, 2, \cdots, N(t)\} \tag{3.2}$$

$$T(t) = \left(\tau_{i,j}(t)\right) \in \mathbb{R}^{N(t)\times N(t)} \tag{3.3}$$

$$\Psi(t) = \{\psi_m(t) | m = 1, 2, \cdots, |\Psi(t)|\} \tag{3.4}$$

For each component, the corresponding individual evolutionary path has been established. Besides, the mathematical notations are abstracted to prescribe the projected deployment of the ATN in each dimension. Each path corresponds to a single projection of the 3D integrated evolution path onto its axis. As a result, the accomplished 3D evolution space is visualized in Figure 27. It is confirmed that only the demand fraction is expressed as a smooth and continuous function, whereas the rest are shown as step functions.

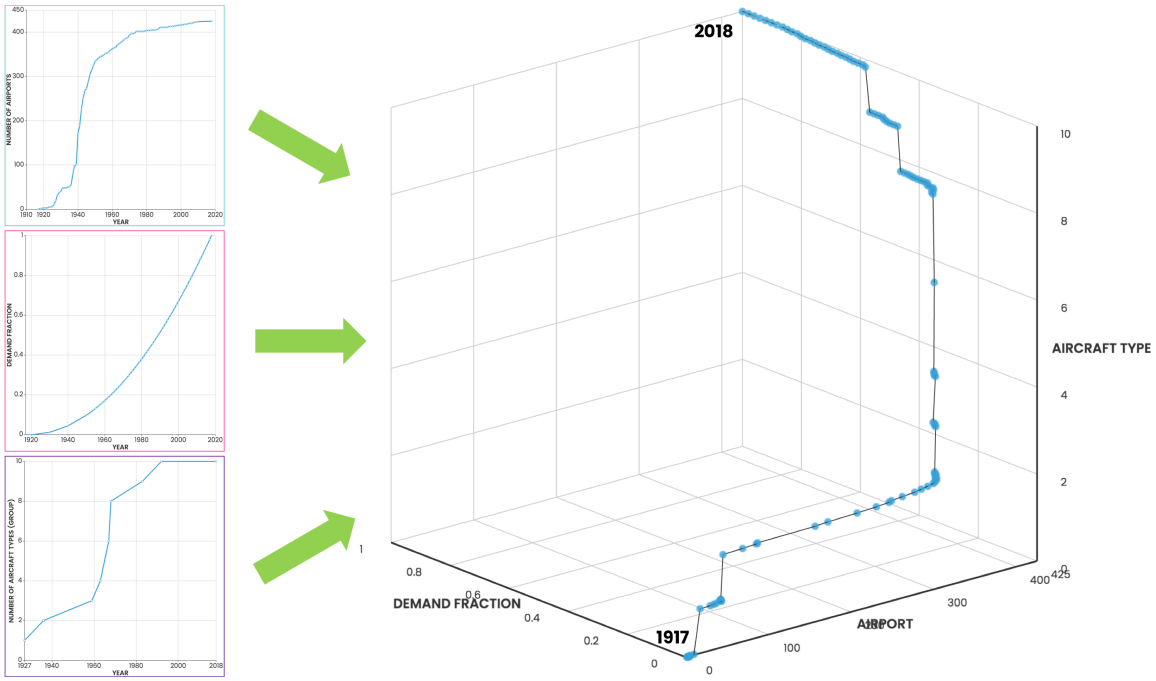


Figure 27 – Complete evolution space of components.

3.5 Result and Discussion

To evaluate the level of realism of the complete evolution space and evolution path, several experimental comparisons are performed. The reference datasets for comparison are retrieved from the research of Yang [110]. Table 13 compares the details of the evolution spaces.

Table 13 – Comparison of the realism of the formulated evolutionary components.

Component	Criterion	This thesis	Yang [110]
Airport	Number of airports	438	53
	History of airports	O	-
	Considered U.S. territory	All regions	CONUS
	Evolution path of airports	Real history	Smooth curve
	Airport capacity	O	-
Demand	Maximum hubs per route	2	1
	Demand percentage	99.9%	~90%
	Demand decomposition	O (two tiers)	-
	Evolution path of demand	Real history	Smooth curve
Aircraft	Number of aircraft types	56	23
	History of aircraft	O	-
	Evolution path of aircraft	Real history	-
	Dynamic operation cost	O	-

Yang's research focused on proving the concept of evolution. As seen in Table 13, this thesis successfully formulated the evolution of components with a significant amount of enhancement in realism. First, this research considers all historical information of all airports in the ATN. Hence, all U.S. territory regions are considered. Second, this research considers up to two-stop routes and 99.9% of the demand is covered. Also, the realistic evolution path of demand is constructed. This was not possible in the previous research, where a smooth curve function is used. Lastly, the evolution of aircraft is firstly accomplished. In Yang's research, the aircraft were abstracted into five representative groups and experience no changes during the simulation. Therefore, the formulation of the evolution of aircraft contributes to augment realism of the architecture model significantly.

In conclusion, the formulated evolution space and evolution path of the components have significantly increased the realism of the architecture model compared to the previous

research. In all components, history was fully considered to model. Moreover, the evolutionary information of this thesis can consider all U.S. territory, 99.9% of demand, and all aircraft types, eventually.

3.6 Chapter Summary

In this chapter, the first research hypothesis is proved by formulating a comprehensive multi-dimensional evolutionary information of the components. First, the airport dataset was acquired from publicly available sources released in two different federal institutions. Then, to describe a variety of airport-related data such as inauguration year, capacity, and geographical properties, six reliable data sources were explored. Hence, the comprehensive evolutionary information for the airports in U.S. territory was successfully constructed. Second, demand and enplanement datasets were cross-compared to extract the complete origin-and-destination information. To that end, the itinerary information in DB1B was explicitly used. However, some challenging hurdles were originated from the lack of data availability; the earliest year of the DB1B was 1993 so that the overall O-D demand and enplanement data before 1993 had to be regressed. The created demand and enplanement datasets followed the real evolutionary path from 1917 to 2018. Third, as for the aircraft types, the historical and performance specifications of the considered aircraft types have been retrieved from institutional references. The most abstruse item was to acquire individual cost information of each aircraft type. The only public source for it was Schedule P-5.2 Form 41 provided by the BTS.

Each sub-research process finally formulated the evolution space, where a variety of different evolution paths could be described. Each dimension is responsible for the

projected deployment of the corresponding evolutionary behavior of the environment. As a result, the realistic evolution path describing the actual deployment of ATN evolution was achieved by integrating the three individual evolution paths altogether. In the long run, the level of augmentation of realism was evaluated by comparing the formulated evolutionary information to that of the previous research. The result showed that the proposed evolutionary information of the components successfully improved realism. Therefore, the first research hypothesis was proved. The developed evolution space can explore a multitude of different evolution paths for abundant what-if scenarios. The actual simulations and explorations will be gradually tackled in all subsequent contents in this thesis.

CHAPTER 4. MODELING MATHEMATICAL RULES OF ATN EVOLUTION

CHAPTER 4 proves the second research hypothesis. At any moment, the surrounding environment urges airlines to respond to it. In the developed architecture model, the airline is assumed to be the representative network architect that makes decisions associated with network segments. There have been, are, and will be a multitude of complex operations among a variety of airlines in the ATN. Complex deployment of strategies for survival will be abstracted, modeled, and finally simulated for the sake of mimicking the emergence of the strong H&S structure. The second research hypothesis is revisited below:

RESEARCH HYPOTHESIS 2

A multi-tiered network evolution approach where an aggregated single airline deploys different policies to construct the primary H&S network and the secondary network can adequately model the rules of ATN evolution.

4.1 Overview

The evolution environment at time t, ENV(t), is defined by $|A(t)|$ airports and $|\Psi(t)|$ aircraft types available with a trip demand prescribed by $\Delta\tau_{i,j}(t)$, the number of passengers who want to travel from α_i to α_j. The interconnection of all elements of ENV(t) draws a curve in a virtual vector space, dubbed 'evolution path.' This abstraction leads to the central hypothesis of this paper as follows: The time-dependent ATN – a tangible form of connectivity, volume, and media – is constructed depending on how policies of the

stakeholders (such as airlines and government) drive and process the network itself at the corresponding evolution point, as shown in the following recursive relationship.

$$\mathrm{ATN}(t+1) = \mathrm{POL}(t, \mathrm{ATN}(t), \mathrm{ENV}(t)), \tag{4.1}$$

where POL stands for airlines' network construction policy (i.e., the dynamics). This is the functional form of the underlying dynamics that interconnects the past and the present as well as the future of the ATN.

In CHAPTER 3, the rules were mentioned in a simplistic statement; it is a function which maps demand to enplanement (see Figure 17). Although the function seems simple, it may look much more difficult as it is highly interactive with components. Even in the literature, there are teeming with research papers that explicitly model the black box from the demand to enplanement in a variety of fields. This thesis challenges against it by striving to pursue the essence of the airline's decision-making mechanism under the given evolutionary circumstances.

4.2 Aggregated Airline Approach for Network Construction

4.2.1 Proposed Approach of Aggregated Airline

This section presents a theoretical foundation of abstracting multiple heterogeneous airlines into one aggregated super-agent. There is no comprehensive dataset or quantitatively known dynamics of the airlines since the information is entirely confidential and proprietary. Even the DB1B, which contains the full itineraries of travels are just a 10% sample from the reporting carriers. It leads to two significant hurdles. First, modeling

the complex dynamics of the airlines' game-theoretic competitions are too elusive to model with a granular fashion. Second, there is little information to formulate the details of strategic policies for each airline enough to define a set of engineering problems.

In addition to these hurdles, the highly unpredictable evolutionary environment exacerbates the elusiveness of the full dynamic modeling about how the airlines let the ATN grow in a full granular level. Therefore, while the established ATN modeling research sought to optimize or enhance the efficiency of the network in a top-level perspective, each airline in the real-world ATN, arguably, instead strives to maximize their profit in a route by route level because not only cannot they have the entire governance on the ATN but also numerous resistances to change the system are already very adamant: regulation, local airport authorities, aircraft manufacturer, etc.

This chapter seeks to capture the essence of airlines' behavioral characteristics; this chapter is devoted to building a concise model that mimics the most general underlying dynamics of airline operations, which are to formulate the H&S structure, eventually. To be more specific, endeavors will not focus on building a model with a granular fashion but instead focus on harnessing the general characteristic behaviors of the ATN evolution. This approach is reasonable since the ATN is neither efficient nor optimized; all network physical resources in the ATN are costly to be swiftly moved around so that they have considerable inertia. Recalling the comprehensive literature review, which shed light on modeling the dynamics of airlines by providing various historical insights, the majority of the ATN modeling studies are based on a premise that there is an ultimate network construction architect who is omnipotent to allocate the fundamental components of the

ATN at his will: aircraft fleet assignment, scheduling, flow control, deciding the number, location, and capacity of hubs.

Embracing this research convention, this thesis applies an aggregated approach for modeling the network construction policy. In a nutshell, an aggregated super-agent airline of which policy is to distribute the demand based on the disutility of selected optimal routes for all O-D pairs. To rationalize the proposed simple abstraction approach with details, this section starts from analyzing the historical data, especially characterizing the evolution of airlines. To sum up this introduction, Figure 28 visualizes the analogy between the real-world airlines and the proposed aggregated airline in this thesis.

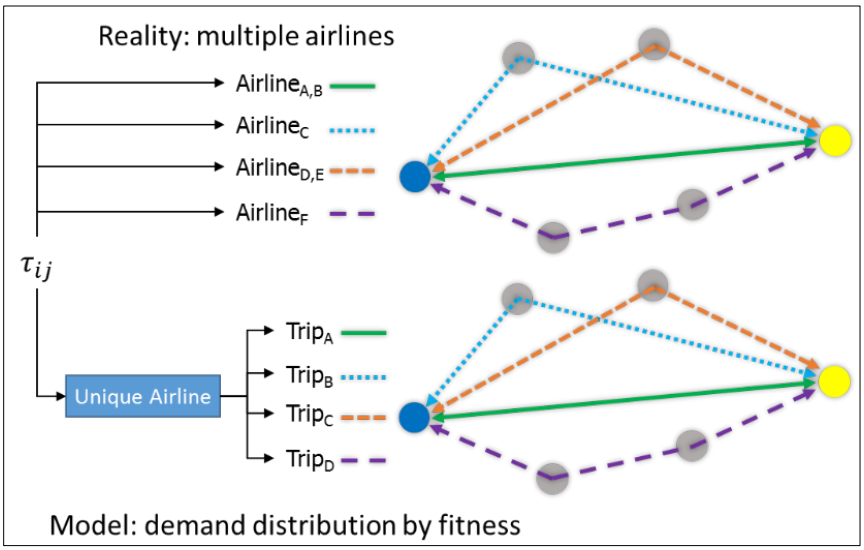


Figure 28 – Notional illustration of an analogy of the proposed aggregated airline approach.

4.2.2 Analogy in Historical Data for Aggregated Airline Approach

According to all available DB1B datasets from 1993 to 2018, the total daily volume of the ATN has increased to 2.16 million. For better understanding, airlines are split into

two categories: major and minor. If an airline has at least 5% of the volume market share, it is classified as major. Otherwise, it is classified as minor. To track the history, the ratio of the direct routes to indirect routes extracted from the DB1B is charted in Figure 29.

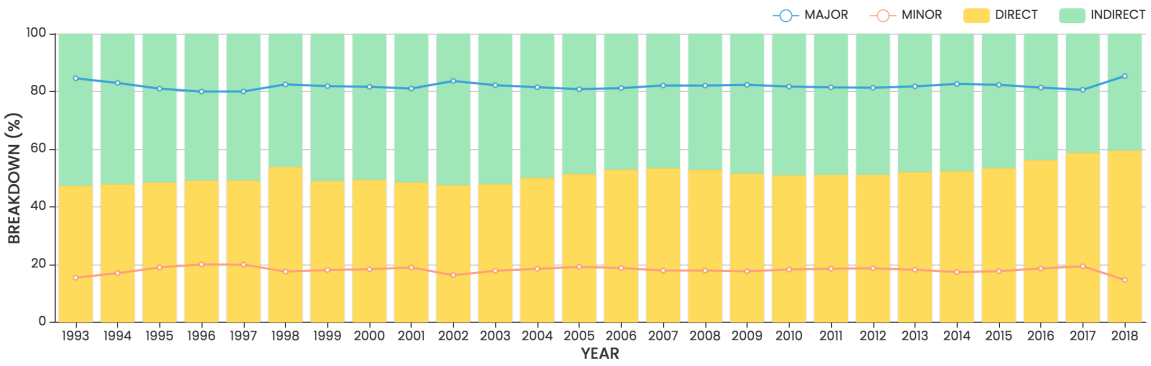


Figure 29 – History of the breakdown of direct/indirect routes.

As seen in the figure, the ratio of direct routes to the indirect routes has varied in a small amount of portion over time. The two lines show the volume breakdown between the major and the minor. It is approximated that the major airlines have served almost 82% of the volume in the ATN for the recent 26 years. Overall, the volume of direct routes keeps slightly increasing. This is an intriguing observation since the increment of the direct route could be construed as the network structural shift towards P2P, or the strong H&S structure is gradually being destructured. Another interpretation is that the number of passengers who wanted to reduce the airtime increases. Figure 30 takes into the deeper level of this data.

Figure 30 shows the history of the major and the minor airlines for the entire 26 years as bar charts. In the figure, the actual number of U.S. airlines is represented as a line chart. While Figure 29 populated every year's data into a 100% relative scale, Figure 30 directly compares year-by-year data by absolute values. In Figure 30, important

information is discovered. In the figure, the number of airlines accounts for those who serve at least one daily passenger for each year. This convention neglects the majority of regional or minute airlines.

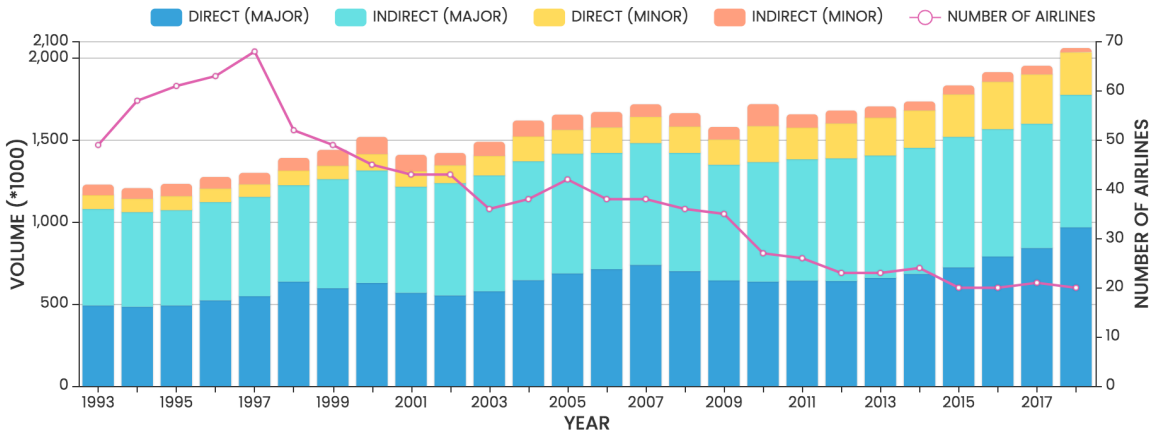


Figure 30 – History of the breakdown of direct/indirect routes for major and minor airlines with the number of airlines from 1993 to 2018.

There are statements to be made from Figure 30. First, the number of airlines has gradually decreased to become 20 in 2018 (line chart.) From this trend, it is glimpsed that M&As have transpired quite actively in a similar manner to the mechanism of H&S structure, in terms of centralization to hub airports, preferential attachment. Trans World Airlines (TWA) and U.S. Airways (US) were merged to American Airlines (AA), Continental Airlines (CA) was merged to United Airlines (UA), Northwest Airlines (NW) was merged to Delta Airlines (DL), and AirTran Airways (FL) was merged to Southwest Airlines (WN); to name several significant M&A events amongst major airlines. At last, the number of major airlines has been reduced from eight in 1993 to four in 2018. In this perspective, the following statement is made:

STATEMENT 1: It is reasonable that modeling the rules of the ATN is mostly about modeling the rules of major airlines.

Second, the volume of minor airlines has been gradually dominated by direct routes. In 2018, 91.01% of total volume are served by direct routes. The rest 9.99% indirect volume corresponds to just 1.23% of total volume; the network created by the minor airlines is almost a perfect pure P2P (i.e., $\forall i, j, \tau_{i,j} = \mathscr{E}_{i,j}$). Thus, it is a reasonable assumption that every traffic by minor airlines is directly served. Thanks to this assumption, an analogy can be made. If the airline tag information of each flight is removed, then the volume of the minor airlines can be absorbed into the direct volume of the major airlines, and it becomes analogous when all accommodated direct routes are distributed by the aggregated one airline, which leads to the following statement:

STATEMENT 2: The network created by minor airlines can be assumed to have a perfect P2P structure.

Third, according to T100D, the number of primary aircraft types (e.g., serving more than 20% of the airline's total volume) of these minor airlines has been one or two at best, as tabulated in Figure 31. In 2018, there are eight airlines classified as the minor: Alaska Airlines (AS), Frontier Airlines (F9), Spirit Airlines (NK), JetBlue Airways (B6), Allegiant Air (G4), Hawaiian Airlines (HA), and Virgin America (VX).

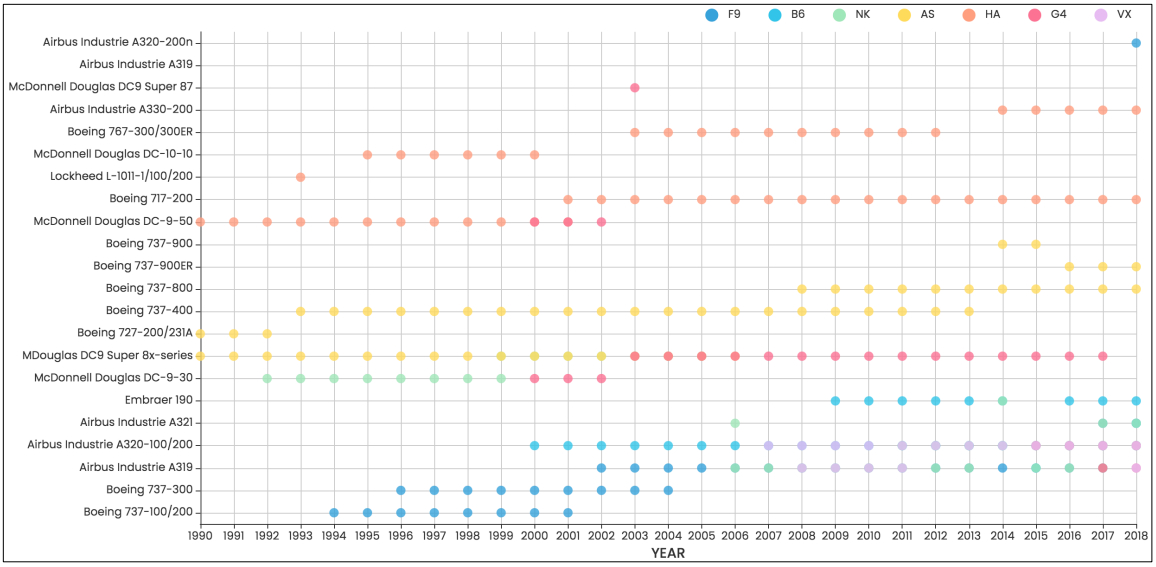


Figure 31 – Aircraft fleet history of minor airlines.

Hence, in the network science's perspective, constructing a P2P-dominated network with one single aircraft type or two similar ones is far simpler than designing the real world's complex ATN. Hence, the network evolution driven by these different minor airlines can be incorporated into the corresponding increment of direct volume for the proposed aggregated super-agent airline in the developed architecture model to make. As a result, the following statement is made:

STATEMENT 3: The number of aircraft types per segment of minor airlines is one or two at best. If it is two, they have similar performance specifications, so it can be assumed to be one aircraft type, too.

Fourth, in more granular detail, one more significant observation can be made. Figure 32 shows the historical variations of the portions of direct routes of the major airlines from 1993 to 2018.

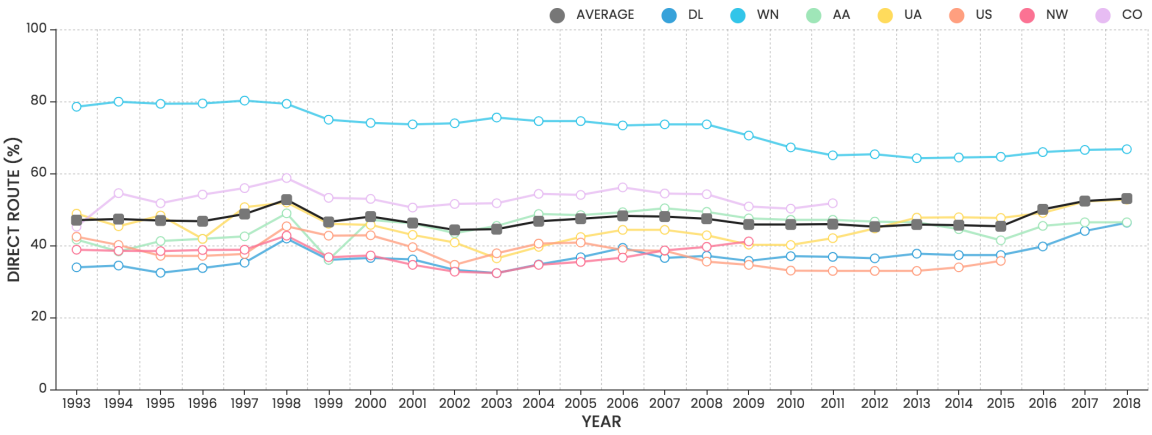


Figure 32 – History of the portion of direct routes of the major airlines.

In Figure 32, the black line is the average of all airlines for each year. Above all, the SouthWest Airlines (WN) is quite distinguished for its generating over 70% of volume via direct routes. This means that WN operates as if it were a minor airline; the network by WN is a P2P-driven network. However, the magnitude of the direct route portion has been gradually shrunken till 2015 and maintained afterward. Note that Northwest Airlines (NW) and Continental Airlines (CO) have been defunct since 2010 and 2012, respectively.

The standard deviation of the line chart for Southwest Airlines (WN) is 5%, whereas those of other line charts are less than 4%. Besides, the gradual reduction of the portion of direct routes of the WN implies that it has slowly transformed its network structure from P2P-driven to H&S-driven for the last 26 years. As for others, the 'average' line chart of all airlines, for the 26 years, shows its average and standard deviation of 47.4% and 2.3%, respectively. Except for the WN, all line charts are lying in a small region of the y-axis: approximately between 35% ~ 50%. As such, the small magnitude of standard deviations reveals that the H&S structure of the ATN has already grasped its maturity in

1993, and even the WN has slowly approached its competitors. This observation yields the following signature statement:

STATEMENT 4: Major airlines have shown a similar trend of route-breakdown composition so that they could be merged into an aggregated metric with little variation.

4.3 Multi-Tier Network Evolution Approach

Previous work [110] considers only the top 53 major airports in the CONUS. This approach is reasonable in two perspectives. First, about 85% of the network volume is associated with the top 53 airports. Second, the rest 15% of the demand that does not meet the service viability is handled very differently from how the 85% demand is accommodated. This is an explicit limitation in that it cannot consider the regional demand. From this perspective, to accommodate the regional demand, the methodologies should be not only different from the proposed but also be able to encompass all 438 airports.

To that end, several multi-tiered network construction approaches have been made as preliminary studies by the author of this thesis. [138,139,140,141] These researches are inspired by the literature papers[33,46,45] reviewed in CHAPTER 2 that analyze the ATN by decomposing it into multiple sub-networks. Hence, this thesis proposes a two-tiered network evolution approach that hybridizes two different algorithms for evolving two sub-networks. The ATN is decomposed into two tiers: primary (Tier-P) and secondary (Tier-S). The Tier-P is for the top major airports and demands $\left(\tau_{i,j}^{[P]}\right)$ that meet the service viability whereas the Tier-S is for the rest minor demands $\left(\tau_{i,j}^{[S]}\right)$. The overall procedure of the network construction is performed as follows:

1. Decompose the demand $(\tau_{i,j})$ into primary & secondary by the given service viability (default value: 5).
2. Construct the primary network (Tier-P) for $\tau_{i,j}^{[P]}$.
3. Construct the secondary network (Tier-S) for $\tau_{i,j}^{[S]}$.
4. Recompose the entire network.
5. Repeat 1~4 till the final time step is reached.

Due to the strong H&S structure, the O-D enplanement matrix sorted by the total volume's descending order is very sparse. According to the DB1B dataset, the average number of destinations of the top major airlines (e.g., DL, WN, AA, and UA) is around 100. Therefore, $\tau_{i,j}^{[P]}$ is only considered in the top 100 airports whereas $\tau_{i,j}^{[S]}$ will be considered in all 438 airports. Figure 33 illustrates the proposed concept of the multi-tier network evolution approach.

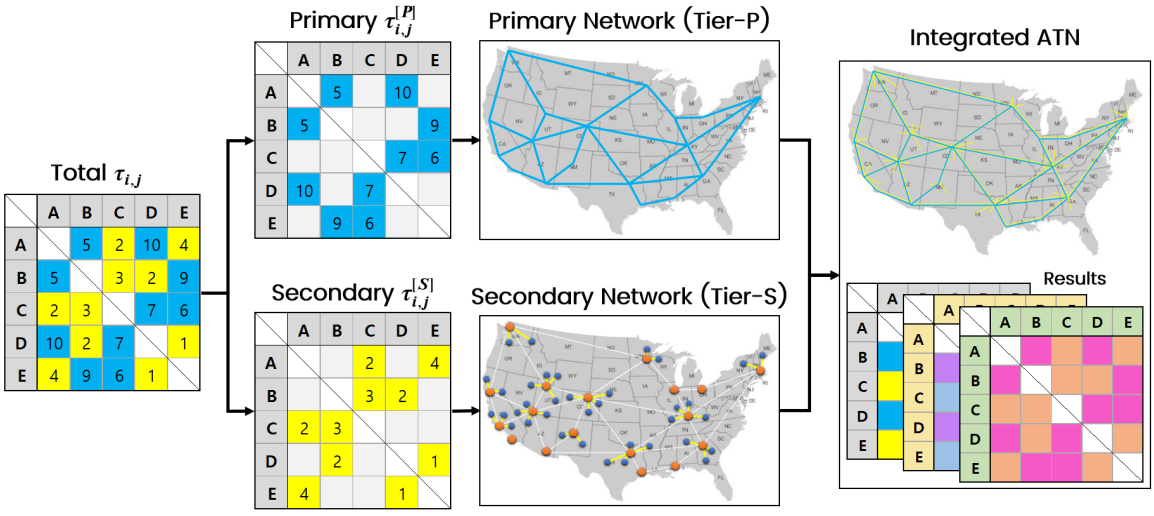


Figure 33 – Notional illustration of the multi-tier network construction approach.

4.4 Construction of Primary Network: Tier-P

4.4.1 Formulation of Disutility

The fundamental architecture proposition drives the construction of both of Tier-P and Tier-S sub-networks; the airline strives to minimize the disutility of operation. Airlines have promptly acknowledged the lucrative nature of harnessing pre-established segments over that of operating direct flights in the service of the passengers. Most attention from the airlines has been concentrated to a few airports, called hubs, in harmony with the strategies of the airlines. Some airports are geographically well-positioned, whereas the others hosted headquarters and maintenance depots of the airlines, influenced by exogenous factors such as socioeconomic status and existing ground transportation infrastructure.

Concerning indirect routes, there are many obvious reasons why airlines take indirect routes (i.e., try to harness their operating bases or hubs as many times as possible) such as cost, efficiency, time, politics, historical inertia, infrastructure. However, investigating these complex factors in a full granular fashion is too overwhelming to be accomplished. In the case of logistics optimization problems, this hurdle is managed by actively adopting a notional coefficient term, which represents the discounted cost per unit flow endowed to incentivize the airlines to use the corresponding hubs. [64]

Embracing that established research, a working hypothesis is introduced: those factors can be encapsulated into a scalar term, dubbed hub discount factor (φh) of airport αh, which reduces the cost of an indirect flight and thus spurs preferential attachment. Then, the airline cost of an O-D operation or a market pair from α_i to α_j can be written as follows:

$$\Delta AlnCost_{i:j} = \begin{cases} \Delta AlnCost_{i,j}, & \text{if direct trip} \\ \left(\Delta AlnCost_{i,h} + \Delta AlnCost_{h,j}\right) \times (1 - \varphi_h), & \text{if 1-stop trip'} \end{cases} \quad (4.2)$$

where the segmental change of airline cost is represented using the following equation. Note that the introduction of $\Delta\tau_{i,j}$.

$$\Delta AlnCost_{i,j} = \min_{y \in \Psi} \Delta AlnCost_{i,j}^{(y)} = \min_{y \in \Psi} \left(AlnCost_{i,j}^{(y)} - AlnCost_{i,j}^{(x)} \right) \quad (4.3)$$

$$= \min_{y \in \Psi} \left(\left\lceil \frac{\mathscr{E}_{i,j} + \Delta\tau_{i,j}}{LF_y \times Cap_y} \right\rceil \times ft_{i,j}^{(y)} \times Cost_y - \left\lceil \frac{\mathscr{E}_{i,j}}{LF_x \times Cap_x} \right\rceil \times ft_{i,j}^{(x)} \times Cost_x \right),$$

where x is the currently assigned aircraft type of segment i-j, LF_y, Cap_y, $Cost_y$ are the load factor, capacity, and hourly cost of aircraft type y, and $ft_{i,j}^{(y)}$ is the segmental flight time for the aircraft type y to fly the segment i-j.

4.4.2 Consideration of Airport Capacity

The airport capacity is an important constraint associated with the operations: departures and arrivals. The airport capacity is the maximum number of allowable operations of an airport per hour. Table 14 tabulates the most recent airport capacity dataset provided by FAA. [142] For simplicity, the capacity is calculated by taking the average ranges of values. Note that the most recent data is of the year 2014. In the table, the statistical data obtained from the airline on-time performance by BTS [143] is also represented as the daily operations. Both values are averaged ones of departures and arrivals. Moreover, Figure 34 plots the hourly distributions of operations of the 30 airports.

Table 14 – Daily operations (2018) & capacity of top 30 major airports (2014).

α	ATL	BOS	BWI	CLT	DCA	DEN	DFW	DTW	EWR	FLL	HNL	IAD	IAH	JFK	LAS
Ops.	1,156	439	314	689	393	807	827	510	451	269	159	267	501	380	476
Cap.	221	121	74	179	71	273	245	181	97	78	119	157	186	89	125
α	LAX	LGA	MCO	MDW	MEM	MIA	MSP	ORD	PHL	PHX	SAN	SEA	SFO	SLC	TPA
Ops.	746	467	392	268	72	246	502	1,154	443	498	289	598	536	349	203
Cap.	172	83	166	76	152	141	162	220	123	142	53	106	105	149	114

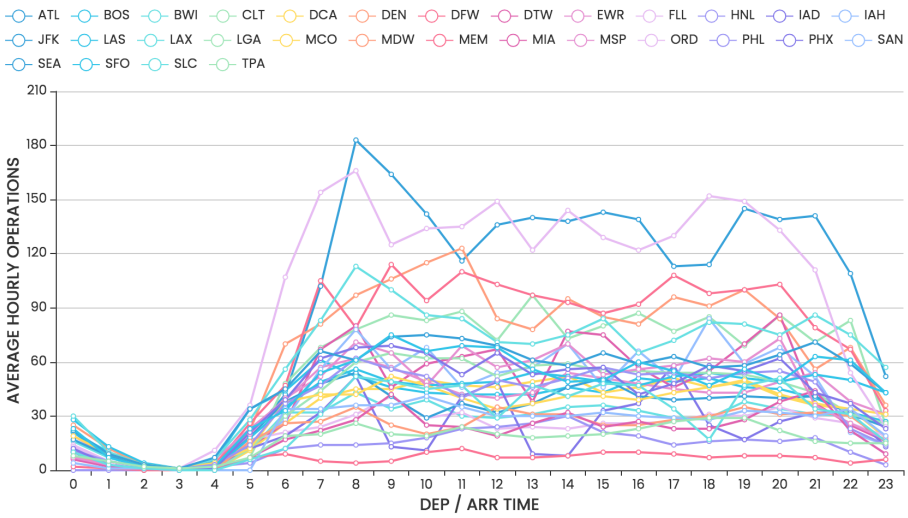


Figure 34 – Average hourly operations of the top 30 airports in July 2018.

According to the figure, the maximum hourly operations are 183 at ATL airport between 8 am and 9 am, followed by ORD airport (166 per hour). In comparison, the operations of each airport in all 24 hours a day in this figure are less than 50% of their corresponding capacity in Table 14. Thus, if the number of operations in the final result from the model is similar to its reference value, then it would be a reasonable assumption that there is a similarity between the operations from the two different sources – real data and model – so that the patterns of hourly operations are also similar.

4.4.3 Construction of Tier-P

The next step is to determine a logistics solution for the given market pair. The simplest way would be to take the minimum disutility among $1 + |A_h|$ routes available and to conclude network segment attributes such as aircraft type and the number of operations. A passenger may choose from multiple flight options that multiple airlines offer. As such, a pluralism should be adopted, but this would cause an excessive computational burden as the evolution progresses. Hence, $|A_h|$ needs to be reduced to a 'reasonable' amount. From a practical perspective, not every airport is useful as a hub. For example, some hubs may have aircraft fly far against the destination, which is unappealing to the airlines and the traveling public alike.

To address these concerns, the authors devised a hub identification algorithm and a probabilistic demand distribution algorithm. The first step is to introduce the traveling public's disutility, dubbed PaxCost. It measures the collective value of passengers' time spent on a trip. Since a type of aircraft is already determined when finding $\Delta AlnCost_{i,j}$, a flight time $\left(ft_{i,j}^{(x)}\right)$ can be computed using its performance data. Then, the change of segmental passenger cost is represented by the following equation:

$$\Delta PaxCost_{i,j}^{(opt)} = \left[PaxCost_{i,j}^{(opt)}\right]_{\text{new}} - \left[PaxCost_{i,j}^{(x)}\right]_{\text{old}} \tag{4.4}$$

$$= \left(\mathscr{E}_{i,j} + \Delta\tau_{i,j}\right) \times ft_{i,j}^{(opt)} \times VoT - \mathscr{E}_{i,j} \times ft_{i,j}^{(x)} \times VoT,$$

where the superscript *opt* is the newly found optimum aircraft type from the airline cost analysis, and *VoT* is the monetary value of time per passenger. For an indirect route, an additional amount (1.5 hours) is added to consider the transfer time in the hub airport to

account for $\Delta PaxCost_{i:j}$. Finally, since all $\Delta AlnCost_{i:j}$ and $\Delta PaxCost_{i:j}$ are ready, the final set of route options A_h^* can be obtained by solving a multi-objective optimization problem, as shown in Figure 35.

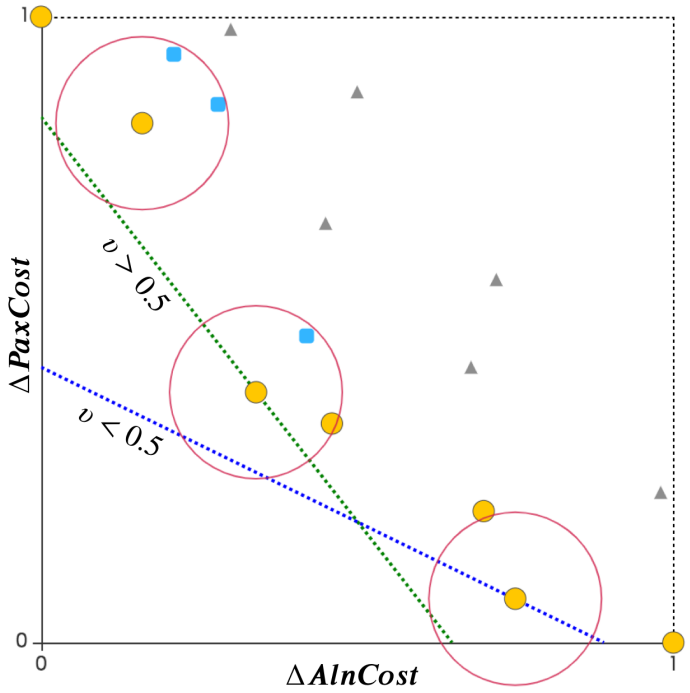


Figure 35 – Pareto and quasi-Pareto optimal routes for demand allocation.

The yellow circles, the three blue rounded squares (within the red boundaries of quasi-Pareto), and the arrows denote the Pareto optima, the quasi-Pareto optima, and the dominated solutions, respectively. Only the Pareto and the quasi-Pareto optima are chosen. In Figure 35, 10 route options (7 Pareto optima + 3 quasi-Pareto optima) are selected, and there are one direct route (obviously unique) and nine indirect routes. Note that the rightmost Pareto-optimum point is the direct route because it has the minimum flight time. The other nine route options will have their route sequences, passing through different hubs.

The critical task is determining how to distribute the passengers through $1 + \left|A_h^*\right|$ routes. It should be noted that a group of passengers is essentially sentient to the ticket price and the schedule, among other factors. As such, a hypothesis that collective sentience of the airlines and the passengers allocates demand depending on a 'fitness' value of each route is adopted. To compute the fitness, the disutility of each route is defined by integrating $\Delta PaxCost_{i:j}$ to $\Delta AlnCost_{i:j}$ as follows:

$$\Delta DU_{i,j} = \gamma \times \left(\upsilon \times \Delta AlnCost_{i,j} + (1-\upsilon) \times \Delta PaxCost_{i,j}\right) \tag{4.5}$$

where γ is a parameter to adjust the impact of disutility on preference and $\upsilon \in [0, 1]$ is a network-tuning parameter; the higher the υ, the stronger an H&S system becomes. Finally, the fitness of route k is calculated by using a logit model (softmax function) as follows:

$$F_{i:j}^k = \frac{e^{-\Delta DU_{i:j}^k}}{\sum_{k=0}^{\left|A_h^*\right|} e^{-\Delta DU_{i:j}^k}}, \tag{4.6}$$

where $\Delta DU_{i:j}^k$ denotes the route disutility using a specific hub airport α_k (α_0 is defined as α_j, i.e., the direct flight). Then, the enplanement of segment i-k is calculated by $\Delta \mathscr{E}_{i,k} = F_{i:j}^k \times \Delta\tau_{i,j}$. If $k > 0$ (i.e., indirect routes), the passengers must transfer at α_k, creating the same enplanement on segments i - k and k - j such that $\Delta \mathscr{E}_{i,k} = \Delta \mathscr{E}_{k,j} = F_{i:j}^k \times \Delta\tau_{i,j}$. Eventually, the total volume contribution created in the network due to $\Delta\tau_{i,j}$ can be computed as follows:

$$\Delta\mathscr{E}_{i:j} = \sum_{k=0}^{|A_h^*|} \Delta\mathscr{E}_{i,k} + \sum_{k=1}^{|A_h^*|} \Delta\mathscr{E}_{k,j} = \Delta\tau_{i,j} + \sum_{k=1}^{|A_h^*|} \Delta\mathscr{E}_{k,j} = \Delta\tau_{i,j} \times \left(1 + \sum_{k=1}^{|A_h^*|} F_{i:j}^{k}\right), \qquad (4.7)$$

where it is straightforward to see $\Delta\mathscr{E}_{i:j}$ in the network is always greater than or equal to $\Delta\tau_{i,j}$ itself as the second term in the parenthesis falls between zero and one. In conclusion, Figure 36 illustrates the entire cell operations of Tier-P network construction.

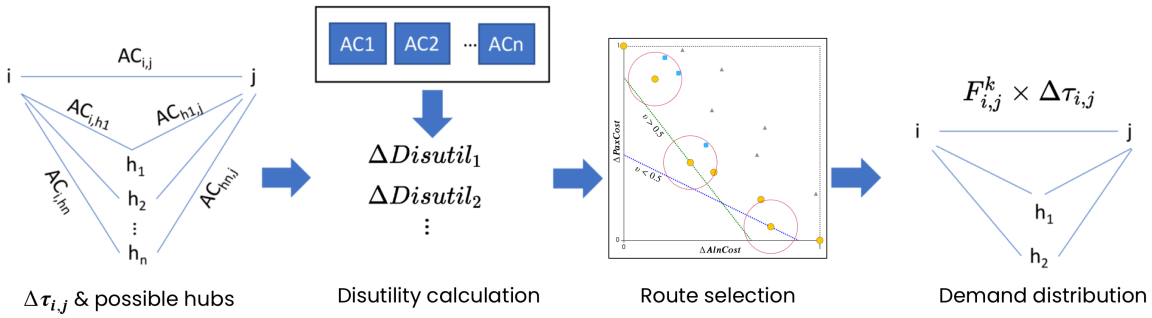


Figure 36 – Notional representation of Tier-P procedures for an arbitrary $\Delta\tau_{i,j}^{[P]}$.

4.5 Construction of Secondary Network: Tier-S

The majority of associate regional or minor airports. Note that the top 100 airports are dubbed primary airports denoted as $A^{[P]}$ and the rest airports are dubbed secondary airports denoted as $A^{[S]}$. Either of origin or destination or both endpoints have shallow values of enplanements. Even though the intense concentration of volume into a handful of major airports can significantly enhance network efficiency, it is challenging for the secondary demand to have direct connections from/to minor airports. Inspired by the author's preliminary researches, [138,139,140,141] this thesis proposes a gravity-inspired network evolution approach. The Tier-S construction process consists of two cell operations: tier-switching from/to the different network tier (i.e., Tier-P → Tier-S and vice versa) and

traveling through the shortest path in the well-established Tier-P within the top 100 airports.

4.5.1 Switching Network Tier by Demand Gravity

All Tier-S trips are wholly encouraged to use the well-established Tier-P network. Hence, the destination of the first segmental flight is identified based on the term 'airport gravity.' The gravity considers two terms: the gravity of airports and the inertia of the travel demand, as represented in the following equation:

$$G_{i:j}^{k} = a \times \frac{\sum_j \mathscr{E}_{k,j}}{{D_{i,j}}^{b}} + c \times \Delta\tau_{i,j}^{[S]} \times \frac{\boldsymbol{d}_{i,j} \bullet \boldsymbol{d}_{i,k}}{\left|\boldsymbol{d}_{i,j}\right|\left|\boldsymbol{d}_{i,k}\right|}, \tag{4.8}$$

where $G_{i:j}^{k}$ is the gravity of airport in the Tier-P (and including), $\boldsymbol{d}_{i,j}$ is the vector from α_i to α_j, and a, b, c are the coefficients to adjust the relative importance of each term. The first term in the right-hand side measures the gravity of whereas the second measures the impact of inertia of the travel demand towards to so that the magnitude can be positive and negative, as shown in $\frac{\boldsymbol{d}_{i,j} \bullet \boldsymbol{d}_{i,k}}{\left|\boldsymbol{d}_{i,j}\right|\left|\boldsymbol{d}_{i,k}\right|}$. It is maximized when $k = j$. Eventually, α_k is chosen by $\alpha_k = \underset{\alpha_l^{[P]} \in A^{[P]}}{argmax}\, G_{i:j}^{l}$. If $k = j$, then the secondary trip becomes a direct trip, and otherwise, the first flight is made to $\alpha_k^{[P]}$, dubbed access airport ($\alpha_k^{[P]} = \alpha_{Acc} \in A^{[P]}$).

4.5.2 Flights via Tier-P Network

The route after the tier-switching is computed by the Dijkstra's algorithm [144] based on the bandwidth: $\frac{\sum_j \mathscr{E}_{k,j}}{D_{i,j}}$. Lastly, if α_j is also not included in the Tier-P network (i.e., $\alpha_j^{[S]}$),

then one more tier-switching occurs from $\alpha_j^{[S]}$. As a result, four types of routes are created: three basic ones from $T^{[S]}$ or a special case from $T^{[P]}$, (i.e., $\alpha_i^{[P]} \rightarrow \alpha_j^{[S]}, \alpha_i^{[S]} \rightarrow \alpha_j^{[P]}, \alpha_i^{[S]} \rightarrow \alpha_j^{[S]}$, and $\alpha_i^{[P]} \rightarrow \alpha_j^{[P]}$). Most Tier-S trips require at least one connection due to the tier-switching. The third type can have up to three layovers (e.g., $\alpha_i^{[S]} \rightarrow \alpha_{Acc,i} \rightarrow \alpha^{[P]} \rightarrow \alpha_{Acc,j} \rightarrow \alpha_j^{[S]}$). Note that if $\alpha_{Acc,i} = \alpha^{[P]} = \alpha_{Acc,j}$ (i.e., $\alpha_i^{[S]}$ and $\alpha_j^{[S]}$ share their access airports), then the number of hubs decreases from three to one.

The Tier-S construction approach has two contributions. First, the emergence of multi-stop routes (i.e., more than one stop) because the Tier-P network comprises only either direct or 1-stop routes. Second, a secondary airport could have multiple connections to its neighbor airports, depending on its destination. This was not possible in previous studies, [138,139,140,141] where only one access airport per secondary airport is allowed. In conclusion, Figure 37 illustrates the entire cell operations of the Tier-S network construction of a Tier-S trip. In the figure, yellow circles and green circles are $\alpha^{[P]}$ and $\alpha^{[S]}$ airports, respectively.

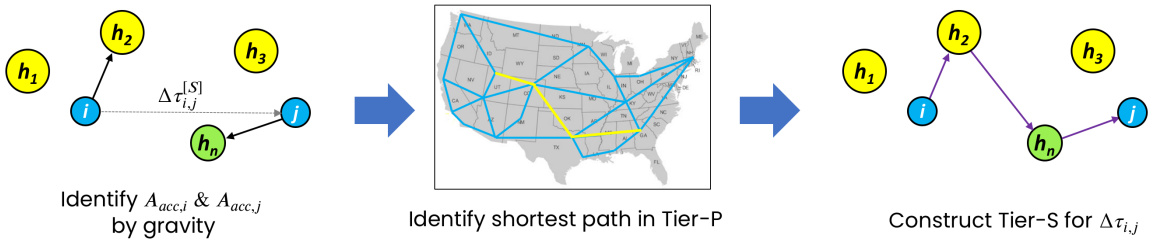


Figure 37 – Notional representation of Tier-S procedures for an arbitrary $\Delta\tau_{i,j}^{[S]}$.

4.6 Complete Architecture Model

4.6.1 Design Parameter

The important design parameters are developed to fulfill the research objective, typically under the goal of realizing the exploratory and interpretative power of the developed architecture model. The important design parameters are organized in Table 15.

Table 15 – Top-level design parameters.

Notation	Name	Type	Range	Default	Description
t	Time step	Sequence	-	-	Evenly discretized sequence of evolution time steps to represent waypoints of years in history.
γ	Discernibility controller	Scalar	$[0, \infty]$	5	Scale controller to make the difference of disutility discernable
υ	Network policy controller	Scalar	[0,1]	0.5	Network policy controller between P2P and H&S structure in disutility calculation
r	Radius of influence	Scalar	[0,1]	0.1	Relative radius for determination of quasi-Pareto options
LF	Limit load factor	Scalar	[0,1]	0.7	Minimum load factor limit for the departure of aircraft
φ	Hub discount factor	Vector	$[-\infty, 1]$	-	Hub factor for hub connection (0: neutral, positive: discount, negative: penalty)
SVF	Service viability factor	Scalar	-	5	Minimum demand for route viability acceptance

Among the design parameters, γ and υ are the most powerful ones. The former can control the discernibility of the disutility so that it can determine the number of Pareto & quasi-Pareto optimum route options. The latter can determine the overall network construction policy of the airline. The radius of influence is also significant, but it mainly affects the number of considered route options for each disutility estimation. Hence, it does

impact on the computational cost. The limit load factor is to evaluate the need for updating the number of operations. Thus, this value will slightly affect the cost, eventually.

4.6.2 Complete Architecture Model

The developed architecture model of this thesis has been explained in a granular fashion. The integrated procedural architecture at a single evolution time step is pictured in Figure 38. Note that every schematic is notional in Figure 38, and the colors of lines are for distinction, not for conveying any quantitative physical implications. This figure describes what happens in just one evolution time step. The entire procedural flow of the developed architecture model is given in Table 16 in the form of pseudo-code.

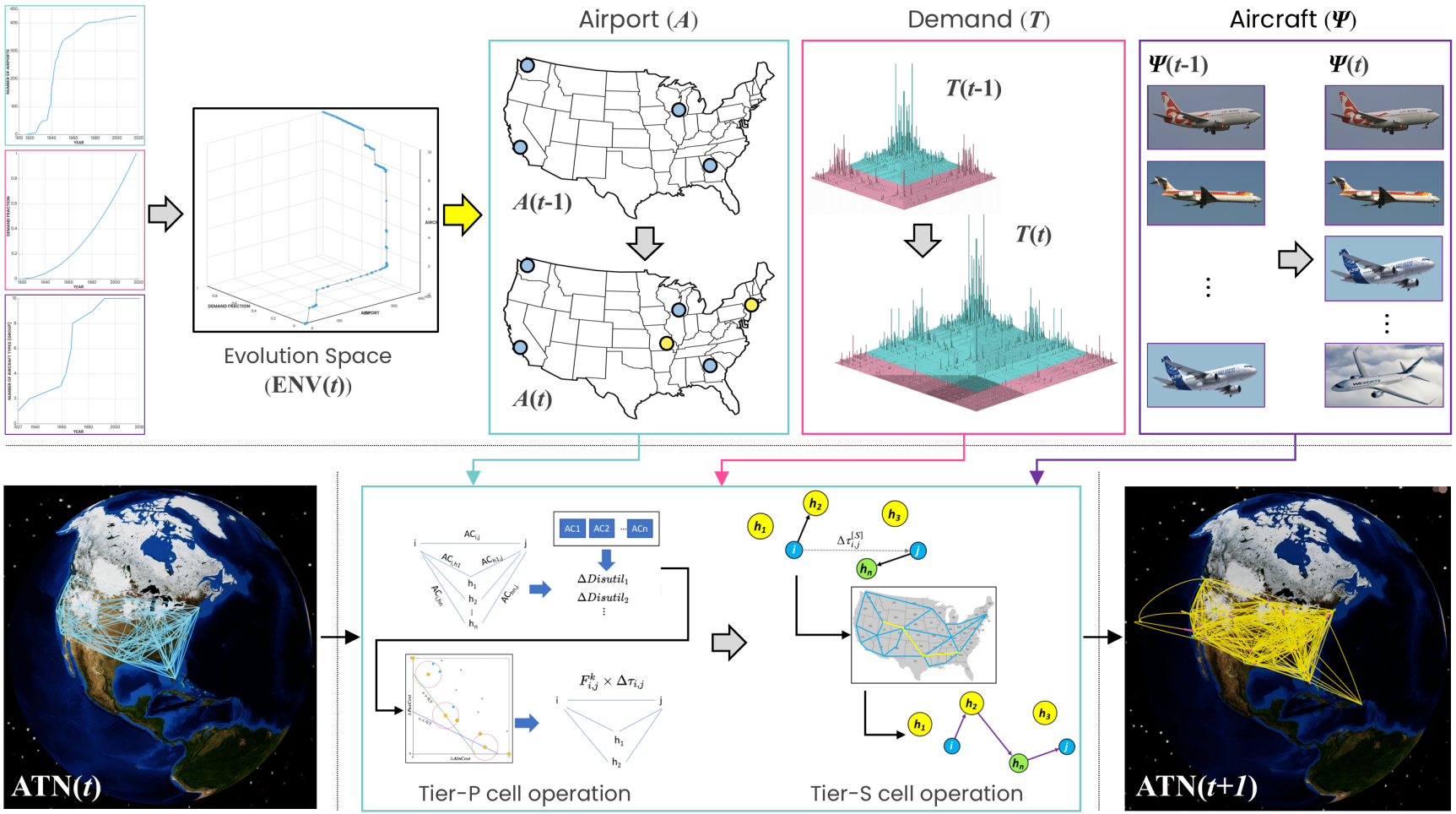


Figure 38 – Integrated design architecture for constructing $ATN(t+1)$ at time step t.

Table 16 – Pseudo-code of the developed architecture model.

$t = 0$; initialize ATN(0)
$t = t + 1$
import ATN(t) and ENV(t) = $\{A(t), T(t), \Psi(t)\}$ from evolution path
for all α_i, α_j: (origin and destination)
  if ($i < j$ and $\Delta\tau_{i,j} >$ SVF):
    set α_i and α_j as active
    direct route $r_{i,l}$: $[\Delta AlnCost_{i:j}, \Delta PaxCost_{i:j}]$ = getCost(i, j, null, $\Delta\tau_{i,j}$, $\Psi(t)$, null)
    for all active hub $\alpha_{h\neq i,j}$:
      1-stop routes $r_{i,h,l}$: $[\Delta AlnCost_{i:j}, \Delta PaxCost_{i:j}]$ = getCost(i, j, h, $\Delta\tau_{i,j}$, $\Psi(t)$, φ_h)
    normalize and populate all $[\Delta AlnCost_{i:j}, \Delta PaxCost_{i:j}]$ into a 2D space
    identify $|A_r|$ Pareto & quasi-Pareto optima and calculate $\Delta DU_{i:j} = \gamma \times \left(\upsilon \times \Delta AlnCost_{i:j} + (1-\upsilon) \times \Delta PaxCost_{i:j}\right)$
    for all $|A_r|$ routes:
      $F_{i:j}^{k} = exp\left(-\Delta DU_{i:j}^{k}\right) / \sum_{l=0}^{|A_r|} exp\left(-\Delta DU_{i:j}^{l}\right)$
      distribute demand to all segments in all routes
  if ($i = j$): $\Delta\mathscr{E}_{i,j} = 0$
for all $\Delta\tau_{i,j}^{[S]} \leq$ SVF: find access airport $\alpha_{Acc,i} \in A_{Acc}$ based on the gravity calculation
for all $\alpha_{Acc,i}, \alpha_{Acc,j}$: find the shortest distance route between $\alpha_{Acc,i}$ and $\alpha_{Acc,j}$ using Dijkstra's algorithm
for all $\Delta\tau_{i,j}^{[S]}$: distribute $\Delta\tau_{i,j}^{[S]}$ considering different cases
  if ($\alpha_i = \alpha_{Acc,i}$ and $\alpha_j = \alpha_{Acc,j}$): $\alpha_i \rightarrow \cdots \rightarrow \alpha_j$ by shortest path
  if ($\alpha_i \neq \alpha_{Acc,i}$ and $\alpha_j = \alpha_{Acc,j}$): $\alpha_i \rightarrow \alpha_{Acc,i} \rightarrow \cdots \rightarrow \alpha_j$ and vice versa
  if ($\alpha_i \neq \alpha_{Acc,i}$ and $\alpha_j \neq \alpha_{Acc,j}$): $\alpha_i \rightarrow \alpha_{Acc,i} \rightarrow \cdots \rightarrow \alpha_{Acc,j} \rightarrow \alpha_j$
finish constructing ATN(t) and save result
if ($t < f$): go to line 2
end algorithm

function getCost($a, b, c, \tau, \Psi, \varphi_c$)
  if ($c = null$):
    $\Delta AlnCost_{a:b} = \min_{\psi_m \in \Psi} \left(\left[AlnCost_{a,b}^{(m)} \right]_{\text{new}} - \left[AlnCost_{a,b}^{(x)} \right]_{\text{old}} \right)$
    $\Delta PaxCost_{a:b} = \left[PaxCost_{a,b}^{(m)} \right]_{\text{new}} - \left[PaxCost_{a,b}^{(x)} \right]_{\text{old}}$
  else:
    $\Delta AlnCost_{a:b} = (1-\varphi_c) \times \left(\Delta AlnCost_{a,c} + \Delta AlnCost_{c,b}\right)$
    $\Delta PaxCost_{a:b} = \left(\Delta PaxCost_{a,c} + \Delta PaxCost_{c,b}\right) + \tau \times ToF \times VoT$
  return $[\Delta AlnCost_{a:b}, \Delta PaxCost_{a:b}]$
end function

4.7 Simulation Framework

The developed architecture model addresses some requirements encompassing different goals listed below:

- High computational performance
- Capability to adopt individual sub-modules for architecture update
- Compatibility with web-based visualization & user-interactivity
- Compatibility with functional programming philosophy

This dissertation uses Google's V8 Javascript as the main programming language for various reasons. For reference, Figure 39 shows a benchmark on different programming languages in 2018. [145]

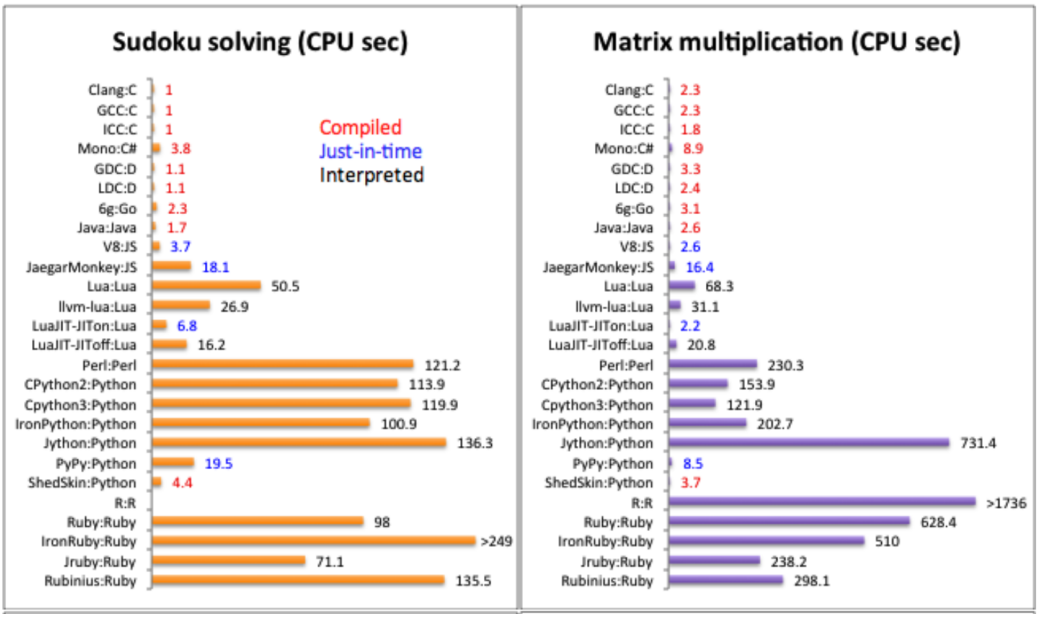


Figure 39 – Performance benchmark on programming languages.

In this performance benchmark, the V8 Javascript is top ranked in all criteria. As such, it was chosen as the programming language for the developed architecture model.

4.7.1 Node.js: *Server-Side Javascript*

Javascript has been the primary language of web programming, mostly front-end programming. Besides, in 2009, there has transpired another revolutionary paradigm shift: the advent of Node.js, the server-side Javascript. Javascript could develop both the front-end and back-end thanks to the introduction of the Node.js.

The Node.js is built on Google Chrome's V8 Javascript engine. Its package manager, dubbed NPM (Node Package Manager), is the largest ecosystem of open-sourced Javascript libraries in the world. Some fundamental properties of the Node.js are listed as follows [146]:

- Performant: Node.js is built on Google Chrome's V8 Javascript engine, so its library is very fast in code execution.
- Asynchronous non-block I/O & event-driven: All APIs of Node.js library is non-blocking. The server moves to the next task immediately after calling a job/function, and notification from its completion helps the server to get a response from the previous call.
- Single-threaded: Node.js follows a single-threaded model with event-looping.

Even though developed for performing back-end development, it can be used as other scientific programming languages, too. Notably, the Non-block I/O & event-driven

feature is speedy and enables the developers to perform numerous different tasks while waiting for the previously submitted jobs are finished, asynchronously.

4.7.2 Functional Programming

Functional Programming (FP) is a programming philosophy where the combinations of functions perform computation. FP treats everything by functions, and even all high-level functions are built by concatenating low-level functions. This philosophy guarantees the data immutability, side-effect-free functions, and declarative paradigm. Considering the adequacy, the developed architecture model can be enhanced by embracing the FP philosophy. For example, different cases could be manifested by combining various elementary functions, which can be repeatedly used to form any kinds of advanced high-level functions. Therefore, the characteristic properties of features in the given scenarios could be easily represented.

4.8 Result and Discussion

One apparent enhancement of realism compared to Yang's research [110] is the formulation of a multi-tiered network evolution approach. In the previous research, only the top 53 major airports were considered. Since these airports covered approximately 90% of the total network volume, it was good enough for proving the conceptual application of evolution in the ATN topology design research. The multiplier approach successfully described the mathematical rules for constructing the network segments due to the research considered only the top 53 major airports.

However, the secondary demand and volume are handled differently because they are not considered economically profitable. As this thesis considers all airports, the same algorithm used in the primary tier network is not relevant. Thus, this thesis developed a multi-tiered network evolution approach to represent the fundamental rules of both primary and secondary tiers. The developed algorithm also allows the minor airports to have multiple connections depending on the relative weight between the gravitational force and inertia of the trip demand. This simple yet relevant approach could successfully let the entire ATN evolve through the simulation. In conclusion, the developed architecture model not only could represent the complex rules of network evolution but also enhanced the realism of the model compared to the previous model from Yang's research. [110]

4.9 Chapter Summary

In this chapter, the second research hypothesis is substantiated. For modeling the complex rules of ATN evolution into its simplistic mathematical representation, an aggregated single airline approach with relevant analogies were identified from the critical observations on the historical datasets. The corresponding four statements are recalled below:

1. It is reasonable that modeling the dynamics of the ATN is mostly about modeling the dynamics of the major airlines.
2. The network created by the minor airlines can be assumed to have a perfect P2P structure.

3. The number of aircraft types per segment of minor airlines is unique or two at best. If it is two, they have similar performance specifications, so it can be assumed to be one aircraft type, too.
4. Major airlines have shown a similar trend of route-breakdown composition so that an aggregated one with little variation can represent their variations.

These statements have been made to substantiate that the network-wise equilibrium state of airline competitions can be simplified to a situation where all tag information of airlines is removed so that the essential mechanism – demand is distributed throughout various routes – is captured.

The second research accomplishment in this chapter is the development of a multi-tiered network evolution approach based on the successful abstraction of the airline. The proposed approach decomposed the network into two sub-network tiers – primary and secondary – to apply proper design architecture for constructing the network and recomposed to create the integrated ATN at each time step, eventually.

Considering the primary network tier, the airline strived to identify the desirable route options in a route-wise level for each O-D trip. All individual O-D demand pairs became able to be transformed into a robust H&S structure network. The well-established primary tier also played a role as the main network tier that interacts with the secondary network tier by allowing the tier-switching process. As for the secondary network construction, the secondary demand was only about 4% of the total demand and mostly non-profitable. The strong sparsity of the demand matrix and many airports were tackled by a simple gravity-inspired algorithm to embody multi-stop routes in the ATN. The

gravity determined the corresponding access airport that the secondary demand should take as the first segmental destination for tier-switching. Then the rest of the trip was handled by the shortest path towards the final destination of the demand.

CHAPTER 5. SIMULATION AND EXPERIMENT

CHAPTER 5 answers the third research question. The first and second research questions have been answered by the research performed in CHAPTER 3 and CHAPTER 4, respectively. Lastly, the developed architecture model must be validated to demonstrate that it can well-follows the reference ATN extensively in a wide variety of criteria. The third research hypothesis is revisited below:

RESEARCH HYPOTHESIS 3

Employing a set of representative metrics and methods from different networks to which the ATN belongs can test the validity of the simulated ATN. Considered network types are general graph, complex network, transportation network, and air transportation network.

The form and function of the ATN are evaluated by various metrics that were deeply investigated in this chapter. Different network metrics capture and characterize different aspects of the function of the ATN. In this perspective, numerous established ATN analysis literature studies have studied the ATN by using many metrics and schemes. Inspired by the observations, this chapter starts with proposing a hierarchical verification and validation (V&V) scheme to perform an organized task to test the third research hypothesis, followed by several research experiments that bolster substantiating the first and second research hypotheses.

5.1 Modeling the ATN Evolution: 1917 ~ 2018.

5.1.1 *Hub Discount Factor Optimization*

To model the ATN evolution from 1917 to 2018, the optimal set of hub discount factor values should be found for the V&V process. The direct implementation of the hub discount factor is the change of airline cost equation, recalled as follows:

$$\Delta AlnCost_{i:j} = \begin{cases} \Delta AlnCost_{i,j}, & \text{if direct trip} \\ \left(\Delta AlnCost_{i,h} + \Delta AlnCost_{h,j}\right) \times (1 - \varphi_h), & \text{if 1-stop trip} \end{cases} \quad (4.2)$$

Only a handful of major airports are given a non-zero hub discount factor. To properly choose the airports, two criteria are made: airlines' operating bases & important airports[124,125,126,127,128,129,130,132,133,134,137] and airports that are ranked top by the generated demand $\left(\sum_j \tau_{i,j}\right)$. [7,8,111] Table 17 tabulates the final set of the selected airports. Let the set of the hub discount factors of these airports be denoted as A_ϕ.

Table 17 – Selected airports for applying hub discount factors (A_ϕ).

ATL	ORD	DEN	LAX	DFW	PHX	MSP	LAS	SFO	SEA	CLT
IAH	DTW	DCA	PHL	SLC	MIA	MDW	BNA	STL	DAL	

The optimization problem is defined as follows:

$$\text{minimize} \sqrt{\sum_i \left(\mathscr{E}_{i,\Sigma}^{Ref} - \mathscr{E}_{i,\Sigma}\right)^2} \quad (5.1)$$

$$\text{subject to } -0.5 \le \varphi_i \le 0.5$$

Here, φ_i is i-th hub discount factor and $\mathscr{E}_{i,\Sigma}^{Ref}$ is of the reference data (T-100D). For performing the optimization, the standard particle swarm optimization algorithm (PSO) provided by ModelCenter®, a famous engineering design/optimization tool, is used. Figure 40 notionally illustrates the defined optimization problem.

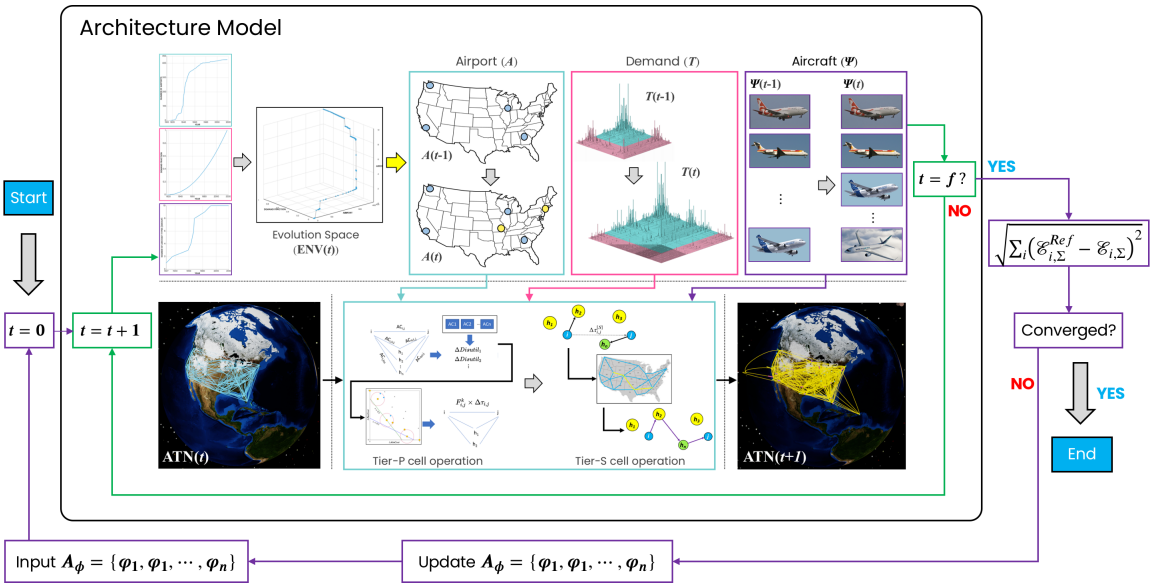


Figure 40 – Notional illustration of the hub discount factor optimization.

Table 18 summarizes the result of the optimum configuration of the hub discount values of A_ϕ^{opt}. The minimized objective function is $\sqrt{\sum_i \left(\mathscr{E}_{i,\Sigma}^{Ref} - \mathscr{E}_{i,\Sigma}\right)^2}_{\min} = 10{,}870$.

Table 18 – Hub discount factor optimization ϕ_h^{opt} result.

α_h	ATL	ORD	DEN	LAX	DFW	PHX	MSP	LAS	SFO	SEA	CLT
ϕ_h^{opt}	0.3241	0.2581	0.4344	0.2861	0.3656	0.1263	0.0842	-0.5	0.2092	0.5932	0.2051

α_h	IAH	DTW	DCA	PHL	SLC	MIA	MDW	BNA	STL	DAL
ϕ_h^{opt}	0.505	0.1154	-0.1701	0.1306	0.0885	0.3954	-0.1154	-0.2821	-0.1415	-0.1489

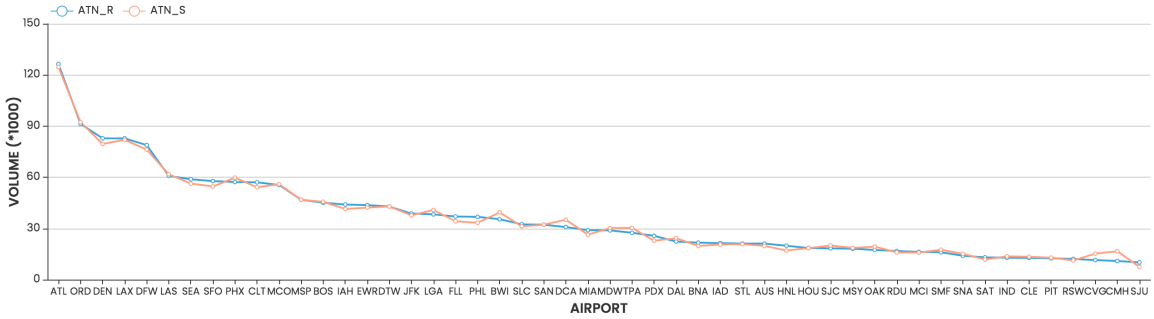


Figure 41 – Volume distribution of the ATN_R and ATN_S.

5.2 Verification and Validation

In the top-level, the ATN is essentially a graph so that just applying some randomly chosen metrics to check the similarity cannot correctly answer the research question. Considering the hierarchical structure of the air transportation network, it is the smallest 'graph' as illustrated in Figure 42 with the selected representative metrics and characteristics of each network category. At the very outside, the general graph exists. Since all other higher network categories include the air transportation network, an air transportation network must show the representative characteristics of its higher categories to a certain extent. As such, the simulated ATN should have a similarity to its reference in those selected metrics and analyses above.

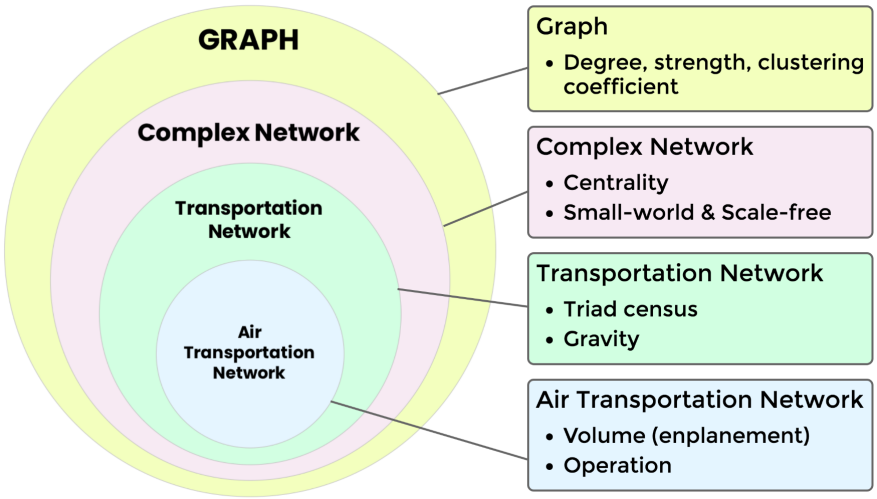


Figure 42 – Hierarchical classification of the air transportation network.

Different metrics and methods that characterize the properties of the air transportation network in different perspectives certainly ease the complexity of the complex air transportation network in their ways, as reviewed in CHAPTER 2. Eventually, 11 representative metrics and properties are selected listed as follows:

- General graph: degree, strength, clustering coefficient
- A complex network: betweenness centrality, closeness centrality, small-world property, scale-free property
- Transportation network: triad census, gravity
- Air transportation network: volume (enplanement), operation

Note that all considered metrics are explained in the appendix. For an arbitrary metric denoted as z, the individual relative error (err_i) is calculated as the following equation:

$$err_i = \left(\sum_i \left(z_i^{Sim} - z_i^{Ref} \right) / \, z_i^{Ref} \right) / N \tag{5.2}$$

As seen, for each airport, the denominator differs. Therefore, this error can be significant as the absolute amount of the metric is nullified. Thus, the global relative error (err_T) is employed as follows:

$$err_T = \sum_i \left(z_i^{Sim} - z_i^{Ref} \right) / \sum_i z_i^{Ref} \tag{5.3}$$

err_T captures the relative error of the total sum of the metric. Thus, the individual significance of airports can be reflected. In a nutshell, err_i and err_T are in a relationship close to that of binary and weighted adjacencies, respectively.

5.2.1 General Graph

5.2.1.1 Node Degree

Figure 43 compares the degree distribution of ATN_S from the developed architecture model and that of ATN_R for the top 50 airports, whereas Table 19 summarizes the errors between them for the top 50 airports.

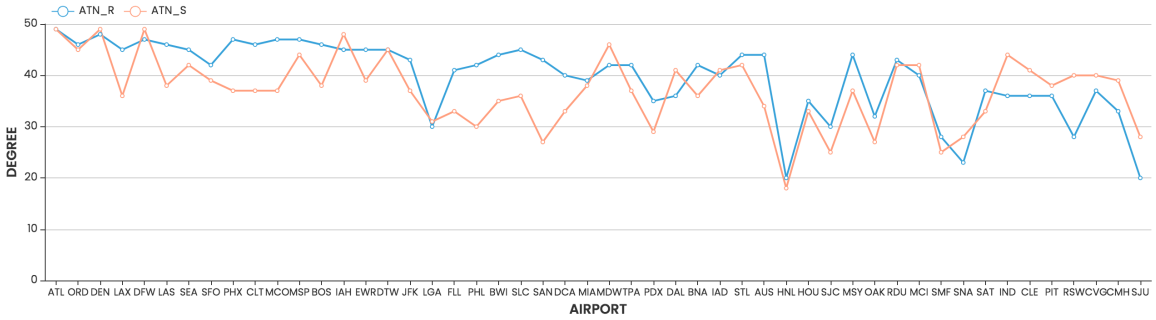


Figure 43 – Degree distribution (top 50 airports).

Table 19 – Comparison of average degree.

	ATN_R	ATN_S
Average $\bar{k}_i$	39.72	37.16
err_i	-	-8.13%
err_T	-	-6.89%

The degree is calculated for the top 50 airports. If there is an error, its amount can be very significant as it is not continuous, discrete. Considering this, the simulated ATN shows a well-matched degree distribution towards the reference ATN. For the top airports, a significant amount of error occurs in PHL, SAN, AUS. However, the average relative error is less than 7%, and the ATN_S still well follows the pattern.

5.2.1.2 Node Strength

Figure 44 compares the strength distribution of ATN_S from the developed architecture model and that of ATN_R for the top 50 airports, whereas Table 20 summarizes the errors between them.

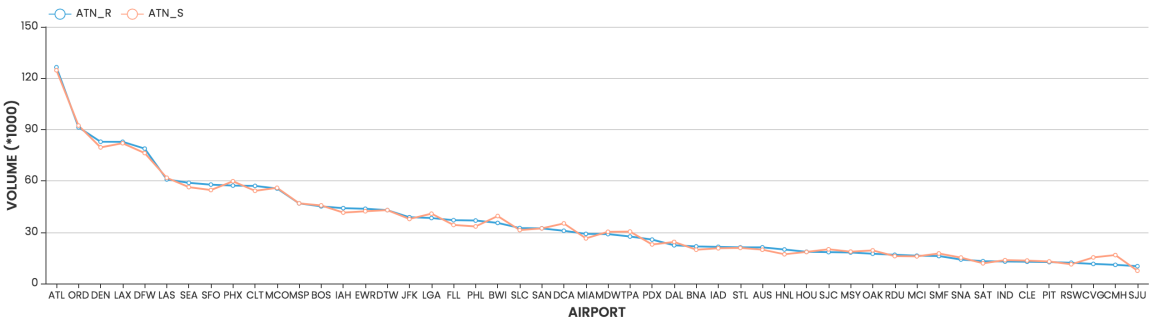


Figure 44 – Strength distribution.

Table 20 – Comparison of strength evaluation.

	ATN_R	ATN_S
err_i	-	-0.27%
err_T	-	-0.44%

According to the figure, the qualitative distribution of the strength is quite like that of the reference ATN. This is because the strength is the variable of the previous optimization problem for modeling the reference ATN. Therefore, for almost every airport, the sums of the enplanement are very accurately mimicking the pattern of the reference ATN. As for the table, the average $\mathfrak{s}_i$ differs by less than 0.5% on average. Both relative errors capture this tiny amount.

5.2.1.3 Clustering Coefficient

Figure 45 shows the clustering coefficient distributions, and Table 21 tabulates the average relative error (%) for both binary & weighted clustering coefficient for the top 50 airports.

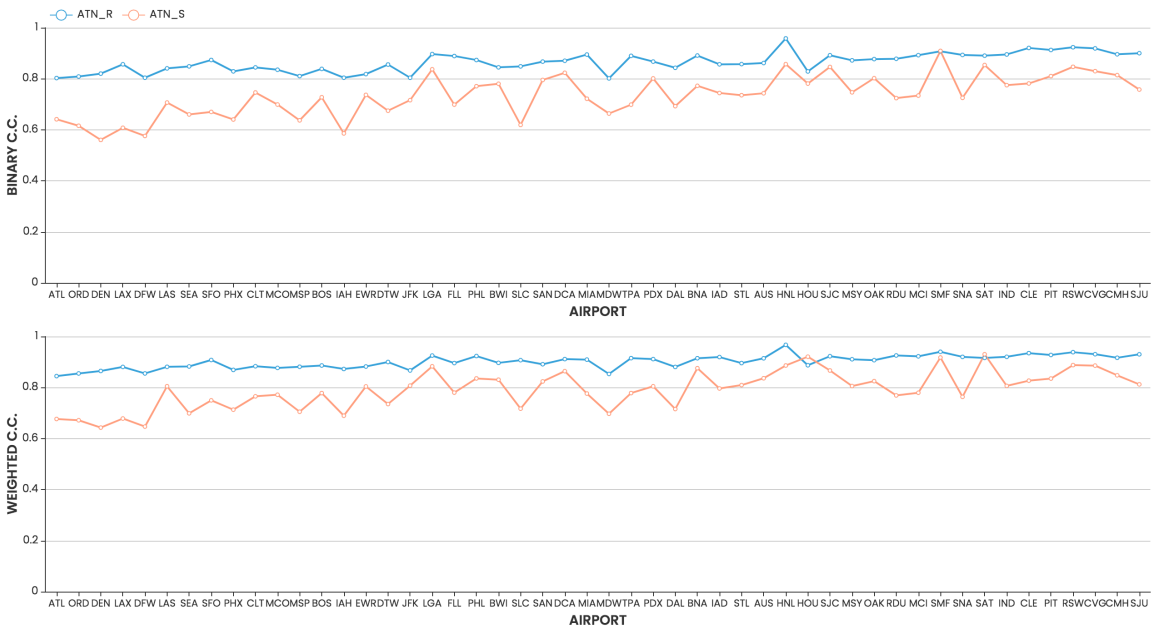


Figure 45 – Binary & weighted distribution of the clustering coefficient.

Table 21 – Comparison of clustering coefficient evaluation.

	Binary (c_i)		Weighted (c_i^w)	
	ATN_R	ATN_S	ATN_R	ATN_S
Average	0.86	0.73	0.90	0.79
err_i	-	-18.80%	-	-14.79%
err_T	-	-17.74%	-	-14.02%

In the figure, it can be confirmed that the overall trend of the distribution of the reference ATN is well-followed by the simulated ATN with an apparent offset in both binary & weighted cases. Therefore, if the ATN_S were slightly shifted up, then the errors would be significantly nullified. The errors are relatively higher than those of other metrics. This is because its mathematical formula is rather complicated, and the range of the absolute value is [0, 1] so that it normalizes the relative significance of each airport for every estimation. Considering these characteristics, it is concluded that the ATN_S still captures the overall trend of the clustering coefficient of the ATN_R.

5.2.2 Complex Networks

5.2.2.1 Betweenness Centrality

When the shortest path is based on the binary adjacency, there can be multiple shortest paths as only the number of segments is the path length. However, if a weighted adjacency calculates it, then there is a single shortest path per O-D pair, in most cases. Thus, In the ATN, the top major hub airports will tend to have high values of $\mathscr{C}_B$.

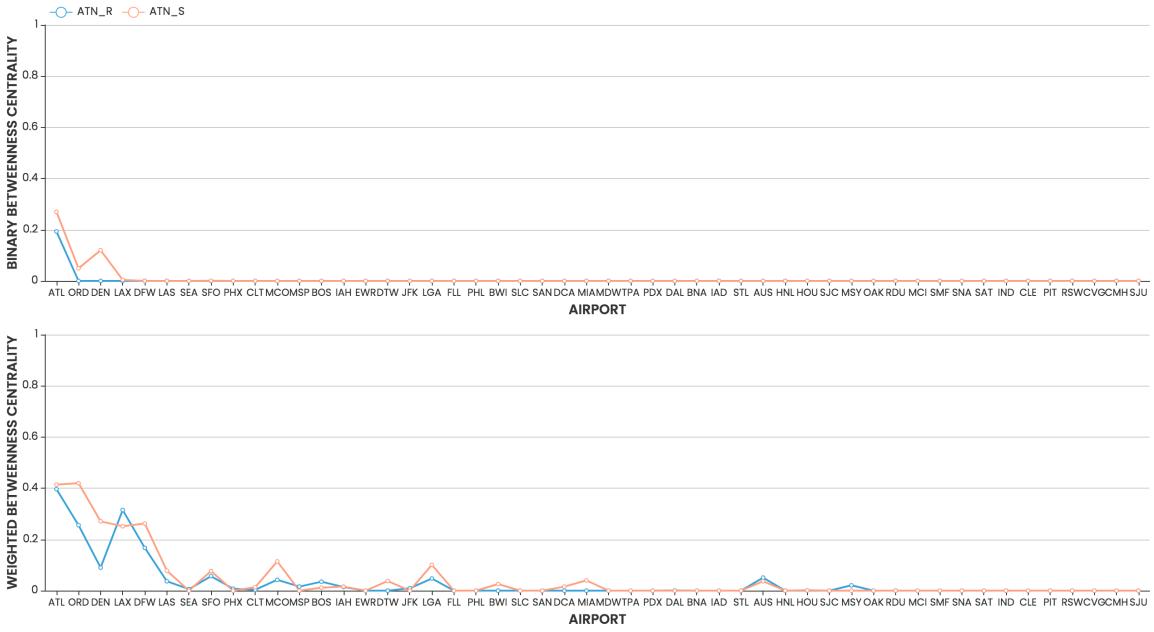


Figure 46 – Betweenness centrality distributions.

Table 22 – Comparison of betweenness centrality evaluation.

	Binary $(\mathscr{C}_{B,i})$		Weighted $(\mathscr{C}^{w}_{B,i})$	
	ATN_R	ATN_S	ATN_R	ATN_S
Average	0.0039	0.0089	0.0314	0.0437
err_i	-	85.68%	-	31.27%
err_T	-	56.47%	-	28.09%

The binary betweenness is concentrated into the ATL airport. Considering the binary betweenness of ATL, $\mathscr{C}^{Sim}_{B,ATL} = 0.0089$ is much larger than $\mathscr{C}^{Ref}_{B,ATL} = 0.0039$ so that the relative error is quite large. As for the weighted betweenness, the ATN_S well matches the behavior of the ATN_R, significantly reducing the relative errors (i.e., almost into half of that of the binary betweenness). Besides, it is clearly observable that the betweenness based on the weighted adjacency shows a much more meaningful and realistic manifestations of the function of the ATN; while the binary one is highly concentrated into a handful of top major airports (i.e., less than 10), the weighted one depicts a relatively more dispersed distribution. In conclusion, the ATN_S can capture the betweenness centrality in the strong H&S structure of the ATN as it shows a quite similar trend to the distribution of the betweenness of the ATN_R.

5.2.2.2 Closeness Centrality

Closeness centrality is increased when regional airports are connected to major airports more closely. Figure 47 and Table 23 summarizes the V&V on closeness centrality.

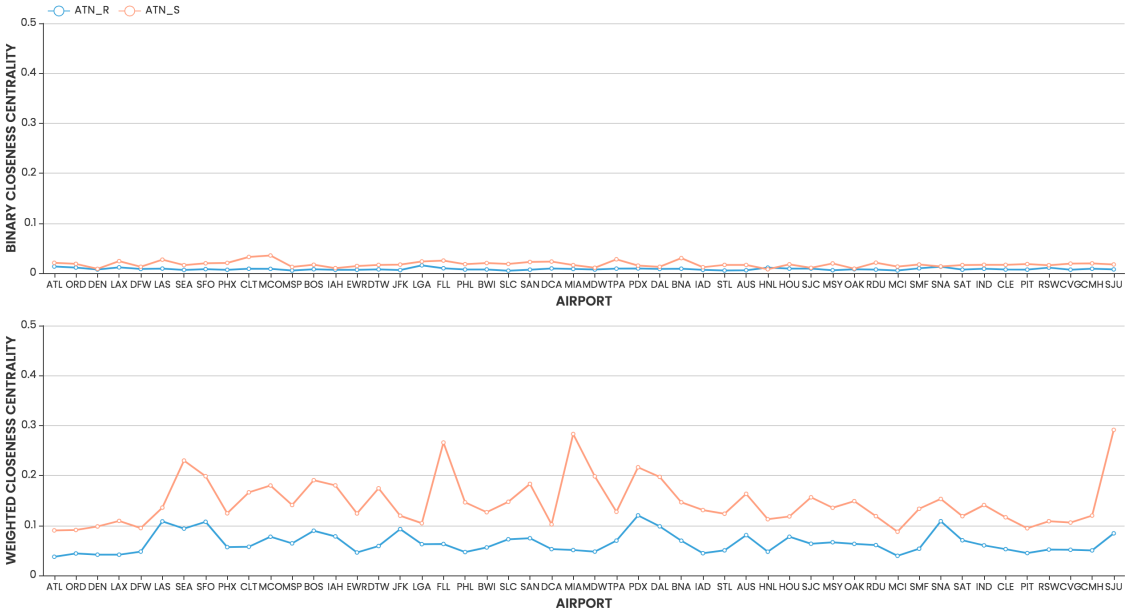


Figure 47 – Closeness centrality distributions.

Table 23 – Comparison of closeness centrality evaluation.

	Binary $\left(\mathscr{C}_{C,i}\right)$		Weighted $\left(\mathscr{C}^{w}_{C,i}\right)$	
	ATN_R	ATN_S	ATN_R	ATN_S
Average	0.0083	0.0180	0.0650	0.1483
err_i	-	49.59%	-	54.29%
err_T	-	54.09%	-	55.95%

Note that the y-axis of the figure is ranged between 0 and 0.5. One easily observable trend is made in the trend of weighted closeness; the weighted closeness of the ATN_S seems to be slightly shifted up from that of the ATN_R. Therefore, the relative errors around 50% can be construed as the offset between the two distributions. Another discussion is from the evenly distributed closeness. Compared to the betweenness, this trend is very different.

According to the closeness centrality distribution, it seems that all airports are similarly crucial in the ATN. However, the average speed of travel and flow from an airport to all others throughout the shortest paths is essential to the minor airports. This is because

indirect routes mostly accommodate their O-D demands via being interconnected to the top major hubs, which may directly travel to most airports in the network. Because of these dynamics, the trend of the closeness centrality in Figure 47 paradoxically emphasizes the significant role of upholding the ATN evolution. Therefore, the more hub airports become essential (i.e., the increment of degree, capacity expansion), the more evenly the closeness distributed. In summary, considering the 'offset' of the result in the weighted closeness, it is concluded that the ATN_S is also validated and verified via the comparison of the closeness.

5.2.2.3 Scale-Free Property

Figure 48 visualizes the cumulative degree distribution of the simulated ATN and the reference ATN. In the figure, the scale-free trend is similar. Note that both axes are logarithmically scaled. Therefore, even though, qualitatively, the trends look analogous, the real errors are quite large. This is because the scale-free property is based on logarithmic values.

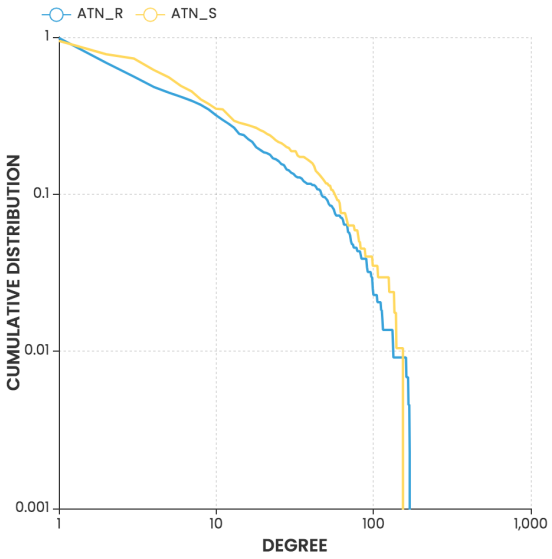
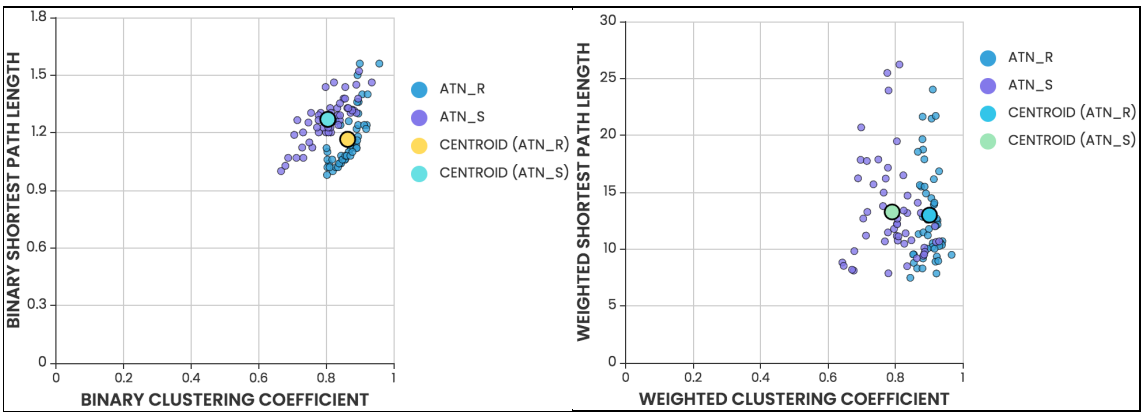


Figure 48 – Scale-free property comparison.

5.2.2.4 Small-World Property

The small-world properties for binary and adjacency representation are visualized in Figure 49. For each scatter plot, the centroids of the simulated and reference ATN are represented. Note that the ordinates for both figures are the average shortest path length.



Binary small-world property Weighted small-world property

Figure 49 – Comparison of small-world properties.

Table 24 – Comparison of small-world property evaluation.

	Binary small-world property		Weighted small-world property	
	ATN_R	ATN_S	ATN_R	ATN_S
Centroid coordinates	(0.86, 1.17)	(0.81, 1.27)	(0.90, 12.99)	(0.79, 13.26)
Distance	-	0.119	-	0.300

The distance is the L2-norm between the centroids. As seen above in the figure, the overall trend of the reference small-world property is quite well-captured by the ATN_S.

First, it turns out that the ATN has a substantial small-world property for both binary and weighted. Second, the average shortest path length of the ATN_S is higher than that of the ATN_R, whereas the distributions of the clustering coefficient are very similar (i.e., the lower and upper bounds are similar). In binary small-world property, the clusters of the two groups separated more than the case of weighted small-world property. As the relative error measurement, the distance is calculated. Compared to the position vectors of the centroids, the distance is small enough to converge to the conclusion that the ATN_S well mimics the small-world property of the ATN_R.

5.2.3 Transportation Network

5.2.3.1 Triad Census

Figure 50 populates the result of the triad census in the simulated and reference ATN. For the 438 airports, there are 13,908,836 triad combinations. Table 25 summarizes the relative errors for each census result.

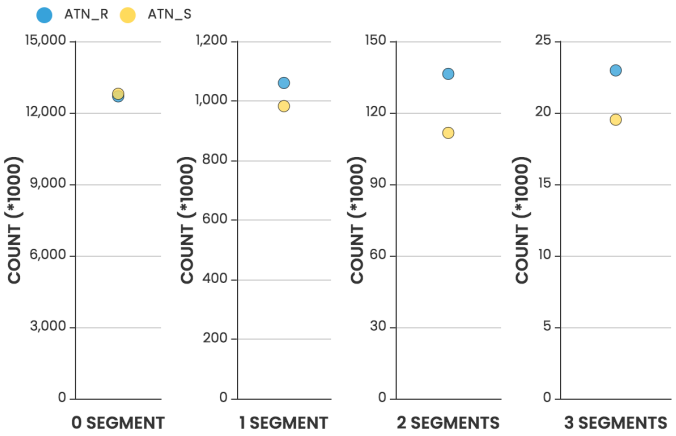


Figure 50 – Comparison of triad census results.

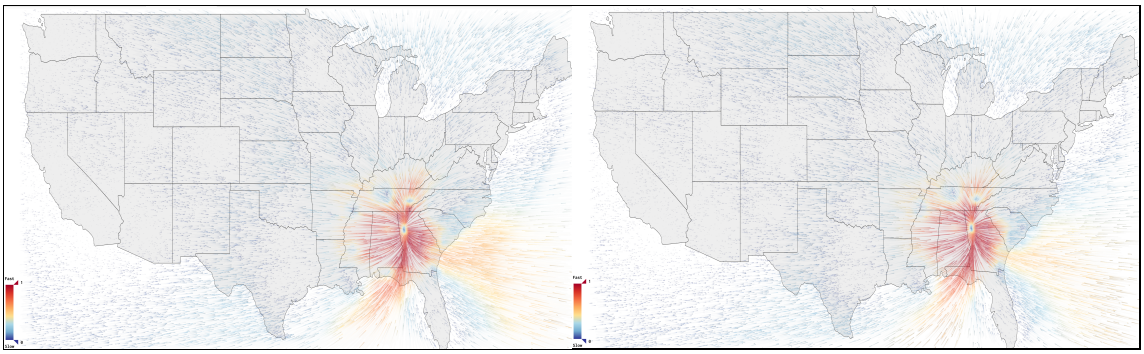
Table 25 – Summary of the triad census results.

Triad type		0-segment	1-segment	2-segment	3-segment
Count	ATN_R	12,689,351	1,060,194	136,327	22,964
	ATN_S	12,795,422	982,339	111,560	19,515
Absolute error		256,071	106,071	-77,855	-24,767
Relative error (%)		2.02	0.84	-7.34	-18.17

In all cases, the triad census estimations are well-matched. The relative error is highest in the 3-segment triad census case, in which all three airports are connected in a triad. However, as the absolute error shows, the 3-segment case is where the smallest difference occurs. In both cases, the most frequent type of triad is the 0-segment type (sets of three airports with no links among them), which shows that the entire network structures are quite sparse. The most barely shown type is the 3-segment type. Therefore, the results of the triad census estimation show the strong H&S structure with a considerable magnitude of the sparsity of the ATN.

5.2.3.2 Gravity Flow Field

The overall flow of gravity is visualized on the CONUS in Figure 51.



Gravity flows of ATN_R Gravity flows of ATN_S

Figure 51 – Visualization of the gravity flow field.

In the figure, all vectors are heading to the region, which is close to the ATL airport that has the most substantial amount of enplanement. In this visualization, both are seldom distinguished, thereby validating that the architecture model yielded the ATN_S that well follows the ATN_R.

Table 26 – Comparison of relative errors of gravity.

	ATN_R	ATN_S
Average $\left\|\underline{\mathscr{G}_x}\right\|$	7.45 pax/mile^2	7.67 pax/mile^2
Relative error	-	2.95%

5.2.4 Air Transportation Network

5.2.4.1 Enplanement (volume)

The passenger enplanement (volume) is the most vivid manifestation of the function of the ATN. This section focuses on the O-D entry-wise enplanement $\left(\mathscr{G}_{i,j}\right)$. Figure 52 represents the 3D visualizations of the enplanement matrices for the reference and simulated ATNs, respectively.

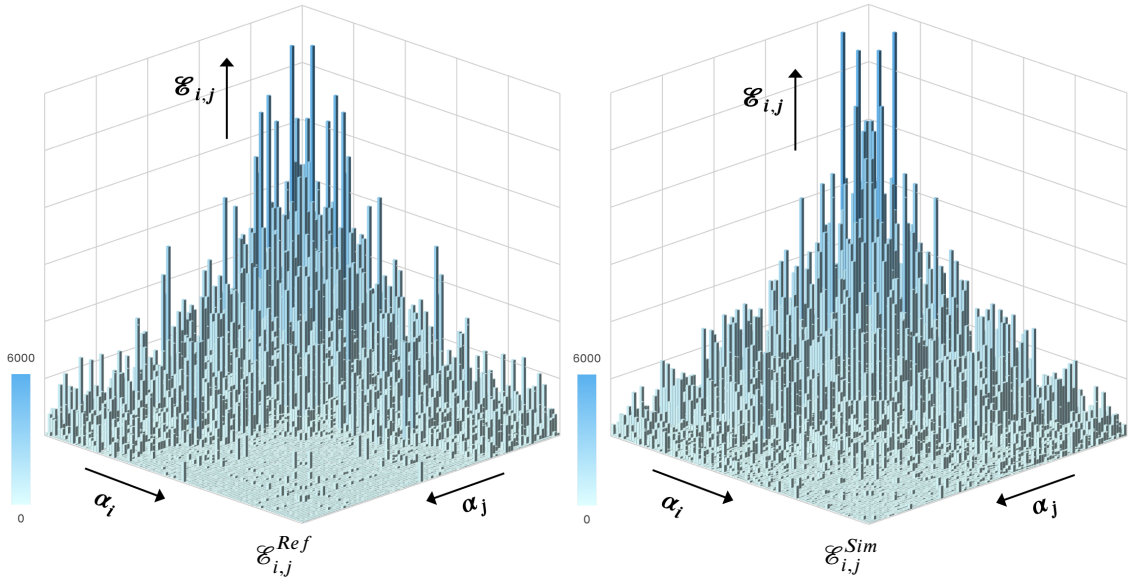


Figure 52 – 3D visualization of enplanement matrices (top 100 airports).

Recalling the objective function in the hub discount optimization uses the row-sum of the enplanements, the entry-wise enplanements do not precisely match to each other. However, the intense concentration in the top major airports is discernibly observed. For example, the maximum enplanements are $\mathscr{E}_{LAX,SFO}^{Ref} = 5{,}450$ and $\mathscr{E}_{LAX,SFO}^{Sim} = 5{,}351$ so that the top segments are relatively well-matched.

As for the error, the maximum absolute error is calculated by $\max\left(\mathscr{E}_{i,j}^{Sim} - \mathscr{E}_{i,j}^{Ref}\right) = 3{,}165$. Its corresponding relative error becomes $\max\left(err_{i,j}\right) = 3{,}165/6{,}052 = 0.5230$ (52.30%). As a result, the heatmap representation of the relative error for the top 50 airports is shown in Figure 53.

In conclusion, the entry-wise enplanements result shows a good agreement of the ATN_S to the ATN_R. In summary, even though there are some segments with relatively high

error (50%), most errors are less than 20%. In conclusion, the comparison of enplanement validates and verifies the model.

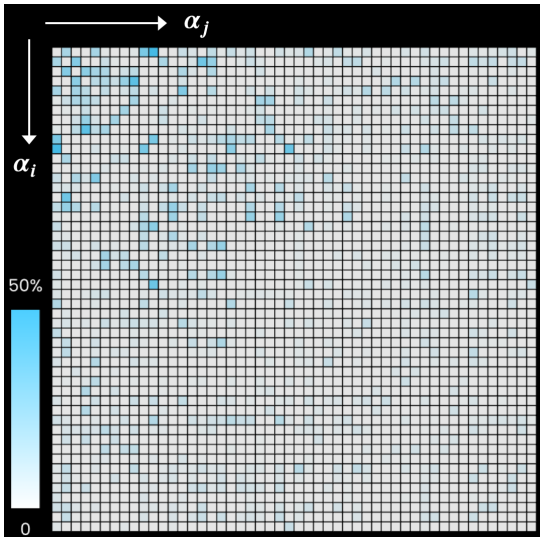


Figure 53 – Heatmap of relative error of enplanements (top 50 airports)

5.2.4.2 Operation

The operation will be investigated for 30 selected airports of which data is publicly given by the FAA. [142] Table 27 summarizes the result.

Table 27 – Comparison of daily operations.

α	ATL	BOS	BWI	CLT	DCA	DEN	DFW	DTW	EWR	FLL	HNL	IAD	IAH	JFK	LAS
Ref.	1,156	439	314	690	393	808	828	510	452	269	159	267	501	380	476
Sim.	1,299	296	277	445	231	567	731	355	279	223	114	139	297	224	471
%	112.4	67.4	88.2	64.5	58.8	70.2	88.3	69.6	61.7	82.9	71.7	52.1	59.3	58.9	98.9
α	LAX	LGA	MCO	MDW	MEM	MIA	MSP	ORD	PHL	PHX	SAN	SEA	SFO	SLC	TPA
Ref.	747	469	392	268	72	246	502	1,155	445	499	289	598	536	349	203
Sim.	683	291	406	207	31	170	458	956	205	497	218	532	456	303	201
%	91.4	62.0	103.6	77.2	43.1	69.1	91.2	82.8	46.1	99.6	75.4	89.0	85.1	86.8	99.0

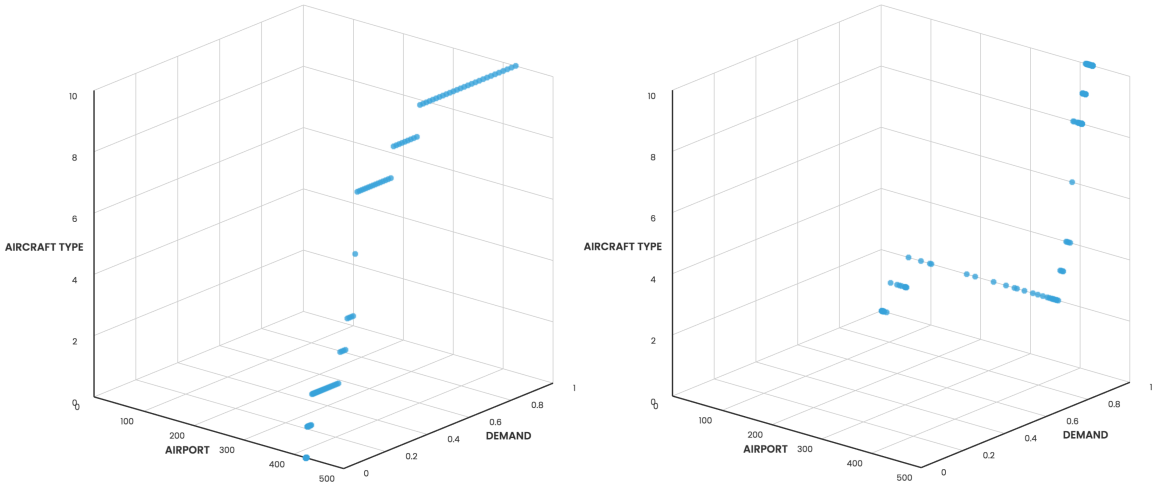
Note that 'Ref.' and 'Sim.' represent the daily operations of ATN_R and ATN_S, respectively. The symbol '%' represents the percentage ratio of operations of the simulation to the reference. From this result, several discussions can be made. First, most airports except for ATL and MCO (i.e., colored in red in the table) show the daily operations less than their reference values, as argued in section 4.4.2. Thus, it can be concluded that the 'hourly operations' have not been violated in any airports, once the hourly distribution can be made following the patterns of the reference; most values are quite like their references. Second, all simulated daily operations are in reasonable amounts compared to the reference. Third, as for the ATL and MCO, their operations are slightly larger than the reference values.

However, those amounts are also reasonable since, recalling Table 14 and Figure 34, the maximum historical operations of ATL and MCO are 183 and 56, respectively, while their capacities are 221, and 166, respectively. In the same way, the maximum hourly operations can be estimated by multiplying the corresponding ratios to the 183 and 56. Eventually, the estimated maximum hourly operations for ATL and MCO become 206 and 58, respectively. As such, both values do not violate the airport constraint. Integrating all these observations and discussions, it is concluded that the ATN_S validates and verifies the ATN_R.

5.3 Experiment 1: Evolution of Airport and Demand

The augmented realism in modeling the evolution can be evaluated by applying different evolution paths of airports and demand. Some unrealistic evolution paths can be explored and compared the model accuracy with the optimized modeling result. Firstly, the spatial expansion of the ATN evolution is assumed to be already completed. Thus, the 438 airports

can be introduced at the first step (i.e., the year 1917). The graph of Figure 54 in the left shows the corresponding evolution path of the number of airports and the integrated evolution path in the evolution space. The second unrealistic path is that of demand; the daily demand is assumed to 100% at the beginning of the simulation. Therefore, airports face a massive saturation issue from its beginning. Hence, this path could be insightful to explore a variety of saturation issues transpiring in different ways. The corresponding 3D path is visualized in Figure 54 on the right.



Unrealistic evolution of airports

Unrealistic evolution of demand

Figure 54 – Evolution paths with unrealistic airports & demand.

Figure 55 illustrates the evolution of the ATNs for different scenarios.

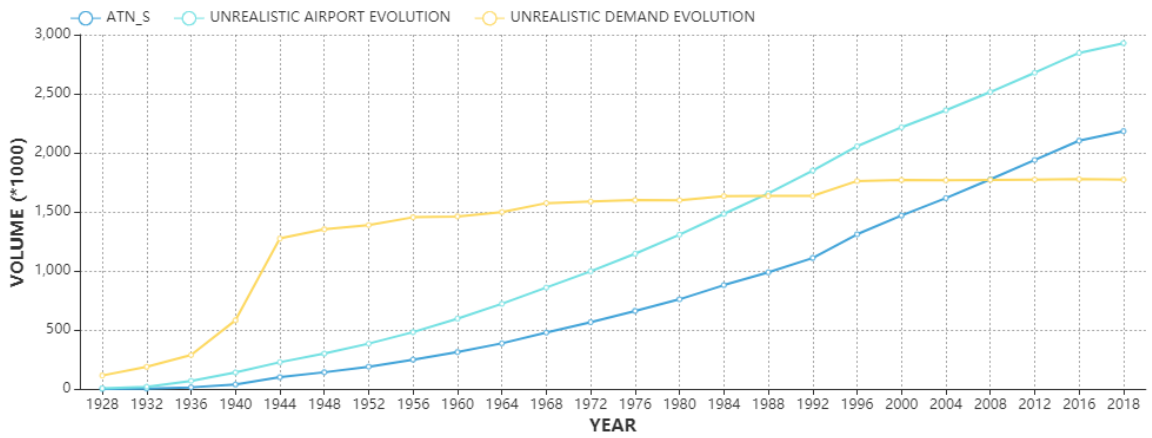


Figure 55 – Experiment results of different unrealistic evolution paths compared with the optimized simulation (ATN_S).

The first observation of the experiment results is that when the airports are fully established, the volume has been explosively increased up to almost 3 million daily volume. It could be construed that the spatially fully expanded ATN can absorb more passengers from other modes, especially for long-range mobility. In the case of the unrealistic evolution of demand, the volume is quite spurred to increase till 1944, in which almost all airports are already established. Therefore, when more airports are introduced, the new demand can create more enplanement. However, since the demand is already saturated (i.e., $\Delta\tau_{i,j}(t > 1) = 0$), the airline finished the network construction, which means that unless there is no airport newly inaugurating, then the current ATN is invariant with time. The slight increment of the volume after 1960 is mainly due to the fleet update by the airline to minimize the disutility in response to the advent of new viable aircraft types. What is accomplished from the experiment by this scenario emphasizes how significant the spatial expansion is for the ATN to evolve chronologically, revealing that establishing an airport from purchasing the land, policy-making, and other complex factors

to reinforce the momentum of airport construction are another underlying drivers of the ATN evolution.

5.4 Experiment 2: Evolution of Aircraft

The enhanced realism of the developed architecture model can be substantiated by conducting an experiment associated with the implementation of the evolution of aircraft. To contrast the simulation accuracy of different configurations, Table 28 tabulates the details about the experiment.

Table 28 – Configurations of experiment cases.

	Yang [110]	Developed architecture model		
		Constant	Limited evolution	Full evolution
Number of aircraft groups	5	10	10	10
Aircraft history	-	-	O	O
Technological evolution	-	-	-	O
Realism	lowest	low	medium	high

As seen above, there is no evolutionary consideration of aircraft technology in Yang's research. Therefore, its relative level of realism is low among the cases. On the contrary, the evolutionary information of aircraft is decomposed into three different cases, each of which has a different level of realism. The term 'aircraft history' mentions whether all aircraft types are viable in all simulation process or the aircraft types are sequentially viable along their market debut years. The term 'technological evolution' represents whether the specifications of aircraft – capacity, range, speed, cost – is static in the simulation or dynamically updated as described in section 3.4.4 (see Table 12.) Note that for an equal comparison, Yang's research is explored with 438 airports, not 53 airports, which is the

original number of the research. The total daily volume of the T-100D in 2018 is 2,160,074. Table 29 compares the results.

Table 29 – Comparison of experiment results.

	Yang [110]	Developed architecture model		
		Constant	Limited evolution	Full evolution
$\sum_{i,j} \mathscr{E}_{i,j}$	1,940,617	2,017,674	2,083,854	2,147,506
Error (absolute)	-219,457	-142,400	-76,220	-12,568
Error (relative, %)	-10.16	-6.59	-3.53	-0.58

The most significant error occurs in Yang's model. The error is partially due to the model's limitation; the model only considers the top 53 major airports with a single network construction approach. Moreover, the most straightforward aircraft information that did not consider any evolution contributes to a large amount of error. In all cases, the results show that the higher the realism is augmented, the more accurate the model becomes. The 'full-evolution' case shows the best accuracy with less than 1% of relative error against the reference volume. Therefore, this experiment demonstrates that the comprehensive evolutionary information of aircraft with augmented realism could better capture the essence of the ATN evolution than the traditional model that focused on proving the evolution concept itself. In conclusion, the result of this experiment also substantiates the first research hypothesis.

5.5 Experiment 3: Artificial Capacity Constraint in ATL

Airport capacity is a critical issue associating with the number of runways, instantaneous weather, precipitation, humidity. According to the FAA's data, the top 30 airports are not saturated in terms of maximum hourly operations. In this experiment, the airport capacity

of ATL is intendedly set to be that of the daily operations of one previous year to explore how the network is changed. A simple formulation of the case study is listed as follows:

- ATL airport is only considered
- The daily capacity is set as the value of 2007
- When the capacity is reached, only connection flights are constrained
- Check the volumetric change between before / after the capacity constraint

The reason why the year was set in 2007 is that there was a massive drop in network volume between 2007 and 2008 due to the economic crisis, as seen in Figure 23. It took several years for the total network volume to exceed that of 2007.

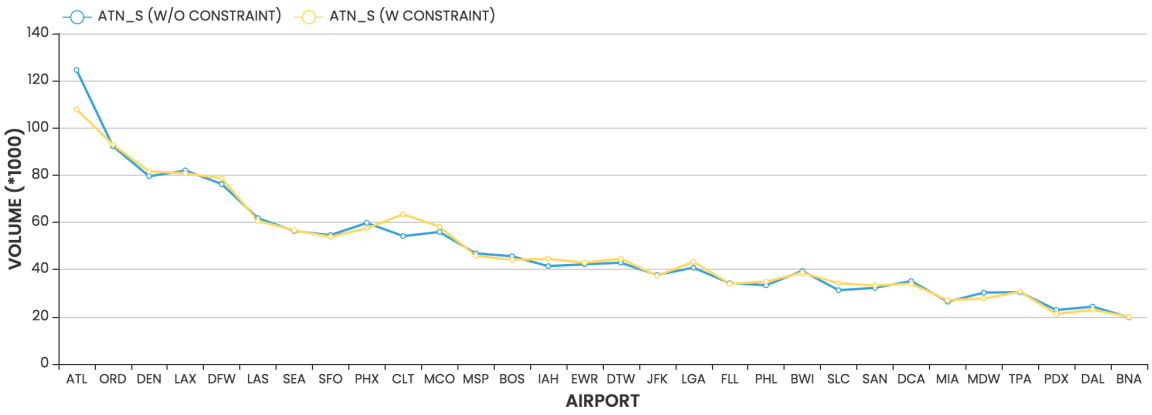


Figure 56 – Capacity constraint case study for ATL airport.

In Figure 56, the maximum daily operations of the ATL was set 1069, which was the average daily operations in 2007. First, due to the capacity constraint, it is confirmed that the volume of ATL has decreased by about 17 thousand, whereas that of CLT has increased by about 9 thousand. Second, even though ATL reached its maximum daily capacity, the total network volume barely changed. Table 30 compares the volumes of several selected airports before and after the capacity constraint in ATL.

Table 30 – Comparison of volumes of capacity constraint case study.

Airport	$\sum_j \mathscr{E}_{i,j}$ (w/o const.)	$\sum_j \mathscr{E}_{i,j}$ (w const.)	Change	%
ATL	124,615	107,889	-16,726	-13.42
CLT	54,191	63,422	+9,231	+17.03
MCO	55,923	58,195	+2,272	+4.06
IAH	41,459	44,557	+3,098	+7.47
LGA	40,821	43,306	+2,485	+6.09
SLC	31,259	34,161	+2,902	+9.28

As seen in the table, when the ATL reaches its capacity, volumes are distributed into other major airports in the vicinity of it. Considering CLT, the airport is located very close to ATL, and it is, in many cases, close to the direct line with ATL for most O-D trips between the west coast to the east coast of the CONUS.

In this simple case study about the capacity constraint, a daily operation capacity was imposed on ATL airport under the assumption that the capacity of ATL was reached in 2007, and no improvement has been made so far. Therefore, from 2008, the network model distributed demands that possibly consider ATL as their hub into other airports, thereby resulting in a significant drop of volume in ATL in 2018, while the neighbor airports received enplanement increment. Therefore, the conducted case study is concluded to be reasonable, albeit the assumptions and approached being simplistic with limitations.

5.6 Chapter Summary

In this chapter, actual modeling of the evolution of the ATN from 1917 to 2018 was performed, and the resultant network topology was validated and verified with a series of network metrics and properties that capture different aspects of the function of the ATN. Then, three experiments focusing on evaluating the level of realism were performed. Some

of the architectures were improved from the previous researches, while this research newly devised others.

As the third research question addresses, validating and verifying the architecture model extensively is crucial to ensure the significant improvement of realism of the simulation. The proposed V&V approach is based on the inference that as the air transportation network is included in its higher network categories, it should show the representative characteristics that evaluate the function, to a certain extent, in those different network perspectives. From the general graph at the top level to the air transportation network, the selected representative metrics and characteristics become less general and more specific toward airline networks. Eventually, the simulated network showed a good agreement in all eleven criteria, demonstrating that it can simulate the evolution of the ATN from 1917 to 2018 with the realistic deployment of evolution.

The performed experiments contrast the level of model's veracity compared to some unrealistic or less realistic evolutions. Each result is supplementary enough to substantiate the first and second research hypotheses. In the first experiment, the unrealistic evolution path of airports and demand showed a significantly diverted deployment of evolution, thereby having many errors. In the second experiment, the evolution of aircraft was explored with a different level of realism in simulation. The results showed that the accuracy has a positive relationship with the level of realism. In the third experiment, an artificial airport capacity was imposed on ATL airport to see the disruption of the network where the most significant airport is operationally saturated. As a result, the network volume of the airport was reduced by about 15%, and other major airports were selected as new hubs during the network evolution.

CHAPTER 6. FORECASTING THE FUTURE ATN DISRUPTED BY SUPERSONIC TRANSPORT

CHAPTER 6 explicitly test the validated and verified exploratory and interpretative capability of the developed architecture model by simulating the future disruption of the ATN caused by the advent of the SuperSonic Transport (SST). As mentioned in CHAPTER 1, various institutes and companies are dedicated to developing the next-generation SST. [21,23,24,25] To that end, a general conceptual SST under the research & development (R&D) will be considered as the upcoming enabler of the supersonic civil air transportation. One thing should be remarked; since all aircraft-related data and information (i.e., cost) is not available, a set of appropriate assumptions will be made. Moreover, the primary market by the SST is not domestic, but international so that this chapter mainly concerns the international SST market.

6.1 Overview

In order to simulate the future ATN evolution, the architecture of the model should be partially modified to make a proper decision-making problem for the airline. Conceivably, the network structure is, highly and mostly, expected to a perfect P2P network. Therefore, several important considerations of architectural change need to be made for a consistent approach:

- A viability factor for the supersonic premium travels is needed.
- Viable SSTs must be able to fly the direct route: $range > D_{i,j}$.

- A certain extent of a buffer region concerning the sonic boom noise should be considered.

6.1.1 Assumption

To begin, the following tuples organizes the scope and assumptions of the simulations in this chapter:

- The demand follows its established evolution path (see Figure 23)
- No new airports will be introduced.
- Established subsonic aircraft types are sufficiently advanced (i.e., no technological evolution after 2018)
- Future SST experiences no technological evolution, either.
- The first-class & business passengers are induced to take the SST. However, they do care about the airfare of the SST flights.
- Acceleration / deceleration of SST is not considered; speed jumps (i.e., 500 mph → 800 mph within 0 second).
- Airports are ready to operate supersonic flights.

These are the minimum set of assumptions so that additional ones on a smaller scale will be made on-demand in corresponding sections.

6.1.2 Service Viability Filter

The service viability of a segment is satisfied when the passengers in the premium cabin reach a good load factor of the SST. To that end, in this thesis, the minimum amount of daily enplanement for an O-D trip to be viable for SST service is set 1,000. In other words,

in order for airlines to make decisions to inaugurate supersonic direct flight service for an O-D, the segmental enplanement must be at least 1,000 such that $\mathscr{E}_{i,j} > 1{,}000$ then 5% of the people (i.e., 50 passengers) are assumed to take the first or business classes. In summary, the enplanement viability denoted as $SVF_{\mathscr{E}} = 1{,}000$. Therefore, the demand matrix for the SST network at time step t $\left(\tau_{i,j}^{SST}(t)\right)$ is defined as follows:

$$\tau_{i,j}^{SST}(t) = \begin{cases} 0.05 \times \tau_{i,j}(t) & \text{if } \mathscr{E}_{i,j}(t) > 1{,}000 \\ 0 & \text{if otherwise} \end{cases} \tag{6.1}$$

As mentioned in the assumption, those 5% portions of each segmental volume are willing to change their trip itinerary from flying via the ticketed established subsonic aircraft type to the SST vehicles. As the chronological progression is deployed, demand monotonically increases, and thus more and more segments will meet the viability factor.

6.1.3 Minimum Travel Time Savings

The value of travel time savings by SSTs must be substantial for all O-D routes. This thesis embraces the corresponding constraint factor dubbed the travel time savings, representing another threshold of travel time savings for a supersonic flight to have to exceed. The constraint for the travel time savings of a segment at time step t $\left(\Delta ft_{i,j}^{SST}(t)\right)$ is defined as the following equation:

$$\Delta ft_{i,j}^{SST}(t) = ft_{i,j}^{(x)}(t) - ft_{i,j}^{SST}(t) > 30 \tag{6.2}$$

Here, $ft_{i,j}^{SST}(t)$ and $ft_{i,j}^{(x)}(t)$ represent the flight time via SST and assigned aircraft type x, respectively, at time step t. In most cases, the distance of the flight trajectory gets longer when it is for SSTs so that the value is positive, meaning the value of travel time savings. Note that the unit is minute. The minimum required value of travel time savings is set 30 minutes.

6.1.4 Sonic Boom Buffer Zone

Another important architecture is the size of the buffer zone concerning the sonic boom. According to the published manual of Concorde, [147,148] there must be a buffer zone to a certain extent for a variety of reasons. It is defined as the offset from actual coastal lines to which SSTs are enforced to fly subsonic from the inland. After SSTs pass the buffer zone to enter an oceanic zone, they can fly supersonic. This is for securing some distance offset to accelerate or decelerate and not cause sonic boom over land. To convey a brief guide for estimating the buffer zone offset, a figure is migrated from reference. [148]

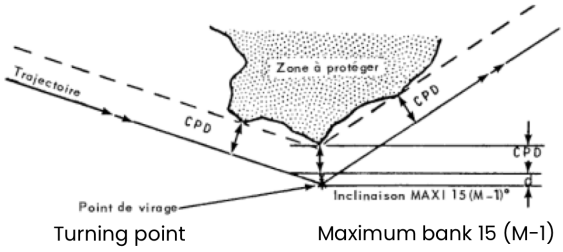


Figure 57 – Buffer zone offset calculation example.

6.1.5 Waypoints for Simplified Shortest Path Finding

To identify the shortest path from the origin to the destination, a simplistic minimal approach is used. To figure out the numerically optimized flight trajectory, this thesis uses

a set of discrete waypoints for finding the optimum shortest path for all O-D SST trips. The waypoints are set on particular locations of the coastal buffer zone so that it is assumed that all SSTs fly towards each waypoint in a straight line, not following the coastal buffer zones.

The entire sphere of the earth will be decomposed into three sub-regions for applying relevant strategies, mostly harnessing the same approach used in the previous domestic SST network construction, which uses a set of discrete waypoints in the vicinity of the continents. Figure 58 shows the geography of the waypoints in for the international SST networks. The red and green circles represent the U.S. waypoints and non-US waypoints, respectively, with the corresponding index.

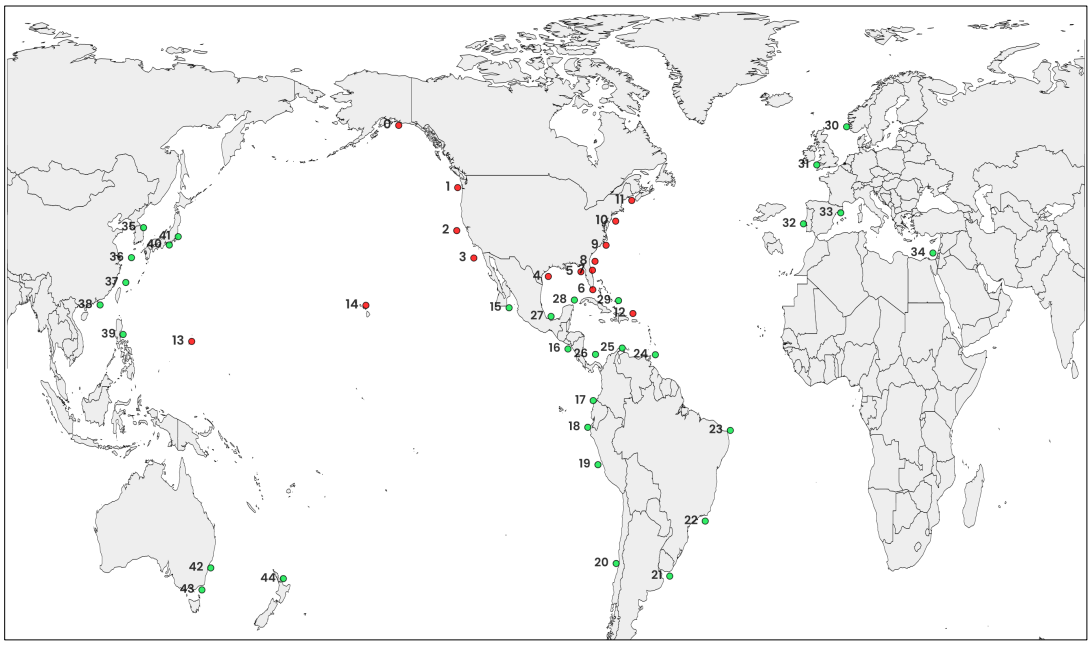


Figure 58 – Waypoints for international SST network.

Including the 15 U.S. waypoints (i.e., 0 ~ 14), there are 30 additional waypoints (i.e., 15 ~ 44) associated with the non-US airports. As mentioned in the previous domestic SST network evolution, a simple rule is abided by for pursuing the simplistic SST flight

trajectory optimization; any flight departed from any airport must fly subsonic till its first waypoint chosen of the shortest path. By the same convention, it is also enforced that the flight trajectory from the final waypoint to the destination must be operated by subsonic.

First, flights between the U.S. and East Asia and Oceania (i.e., waypoints: 35 ~ 44) are mostly above the Pacific Ocean or north pole of the earth, where the supersonic cruise is possible. Second, all direct flights between the U.S. and Europe & Central/West Asia continents (i.e., waypoints: 30 ~ 34) pass through the Atlantic Ocean. Differently from the CONUS and its coastal lines, the geographical composition of these regions is quite complicated so that till the SST enters the Atlantic Ocean, it should cruise in subsonic speed. The five waypoints in this region are defined considering this issue. Note that the only accepted supersonic segment between the green waypoints the segment between point 33 and point 34, which connects IST (Istanbul, Turkey) and BCN (Barcelona, Spain). Third, the flights between airports inside the America continent (i.e., waypoints: 15 ~ 29). These airports are almost in the coastal lines facing either of the Pacific Ocean or the Atlantic Ocean. Therefore, once the SST enters any closest waypoint, then the trip can be supersonically accommodated till the closest waypoint of the destination, thanks to the geographical advantage.

Note that the simplified approach for identifying the shortest path for the SST network by using a set of waypoints is neither accurate nor universal as opposed to the numerical exact airport-to-airport calculation. However, the most benefit on the flight duration is obtained while flying above the Atlantic Ocean so that the first goal for all departures from these regions is to get to the closest waypoint by subsonic; then the

supersonic cruise is initiated, eventually till the final red waypoint (i.e., the waypoint in the U.S.) that is closest to the destination.

6.2 Evolution of Components

6.2.1 Airport & Demand

There are no full O-D itinerary datasets like DB1B by BTS. Recalling Figure 17, in order to result in a simulated network, demand must be provided as one of the evolutionary components. To that end, the reference data is based on the T-100I by BTS. [149] The T-100I is a database that contains international non-stop segment data reported by both air carriers. Like the T-100D database, the T-100I segment has almost the same composition of data.

In order to obtain the itinerary information in an approximated manner, the scheduled flight information in the Google Flights [150] was gathered for the O-D pairs. Figure 59 is an example display of the Google Flights for the ticket from JFK to LHR departing on December 14th and returning on December 21st.

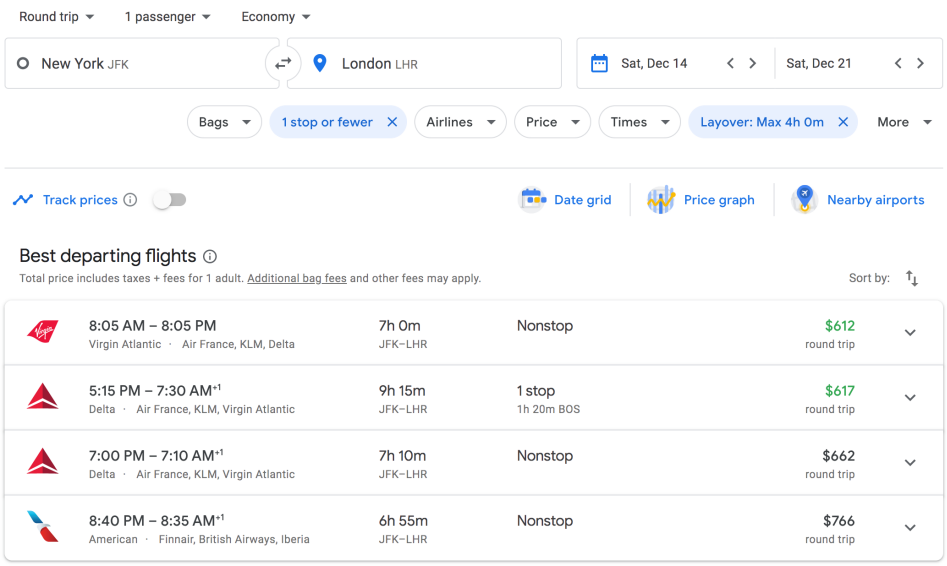


Figure 59 – Google flights example display.

According to the planned schedule, there is one recommended one-stop flight route taking BOS as the hub. There are other indirect route options which are not recommended due to either the high price or the substantially long layover time. In this perspective, assuming there are no multi-stop international travels, the possible hub airports for each O-D pair can be identified, although the planned schedule is not what happened, what could happen in the future, depending on the number of sales of the tickets.

Conceivably, choosing a reasonable amount of layover time in hubs, and a reasonable amount of ticket price, a reasonable number of hub airports for all O-D can be acquired. Top 37 U.S. airports out of 159 and top 83 non-US airports out of 309 cover 97% and 95% of all volume in T-100I, respectively. To reduce the O-D pair itinerary hub list mining, the 37 U.S. airports and 83 non-US airports are chosen. As a result, a total of $37 \times 83 = 3{,}071$ O-D pairs will be explored through Google Flights. To do this, several assumptions are made.

- Only non-stop and 1-stop flights are considered.
- Airports in T-100I are assumed the true origins and destinations.
- The maximum layover time is 4 hours.
- Departure date is 2019 July 20th (Saturday)
- The returning date is 2019 July 27th (Saturday)

Under these assumptions, all O-D pair-wise possible lists of hubs based on Google Flights are gathered. The most reasonable O-D demand datasets are to be found by solving the following optimization problem.

$$\text{minimize} \quad \left| \mathscr{E}_{\Sigma}^{Ref} - \sum\nolimits_{i,j,k} \tau_{i,j,k} \right|$$

$$\text{subject to} \quad \mathscr{E}_{i,j} \geq 0 \ \forall i,j \tag{2}$$

$$\tau_{i,j} \geq 0 \ \forall i,j \tag{2}$$

$$\begin{aligned} \min\left(0.95 \times \mathscr{E}_{i,j}^{Ref}, \mathscr{E}_{i,j}^{Ref} - 100\right) &\leq \sum\nolimits_{k} \left(\tau_{i,j,k} + \tau_{k,i,j}\right) \\ &\leq \max\left(1.05 \times \mathscr{E}_{i,j}^{Ref}, \mathscr{E}_{i,j}^{Ref} + 100\right) \end{aligned} \tag{3}$$

$$\begin{aligned} \min\left(0.95 \times \mathscr{E}_{i,\Sigma}^{Ref}, \mathscr{E}_{i,\Sigma}^{Ref} - 800\right) &\leq \sum\nolimits_{j,k} \left(\tau_{i,j,k} + \tau_{k,i,j}\right) \\ &\leq \max\left(1.05 \times \mathscr{E}_{i,\Sigma}^{Ref}, \mathscr{E}_{i,\Sigma}^{Ref} + 800\right) \end{aligned} \tag{4}$$

Here, $\mathscr{E}_{i,j}^{Ref}$ represents the enplanement of the T-100I from α_i to α_j, and $\tau_{i,j,k}$ represents the partial amount of $\tau_{i,k}$ that distributed to hub j. The indices start from 0, which means the direct (no hub). The objective function is to minimize the total T-100I volume for the selected airports and the sum of all distributed demand throughout all possible hubs

gathered from the Google Flights. Constraint 3 is to entry-wise condition to enforce the absolute amount of the error $\left(\mathscr{E}_{i,j}^{Ref} - \mathscr{E}_{i,j}\right)$ to be within the given range. Constraint 4 is to row-wise condition to enforce the absolute amount of the error $\left(\mathscr{E}_{i,\Sigma}^{Ref} - \mathscr{E}_{i,\Sigma}\right)$ to be within the given range. Note that the number of k, the number of hubs is different for each O-D pair.

Inevitably, this optimization problem contains a certain amount of errors as there is no information or guidance about the relative importance of the hub airports. This problem was solved by using the linear programming feature in Matlab®. Figure 60 shows the result of the converged solutions and corresponding $\tau_{i,\Sigma}$ information.

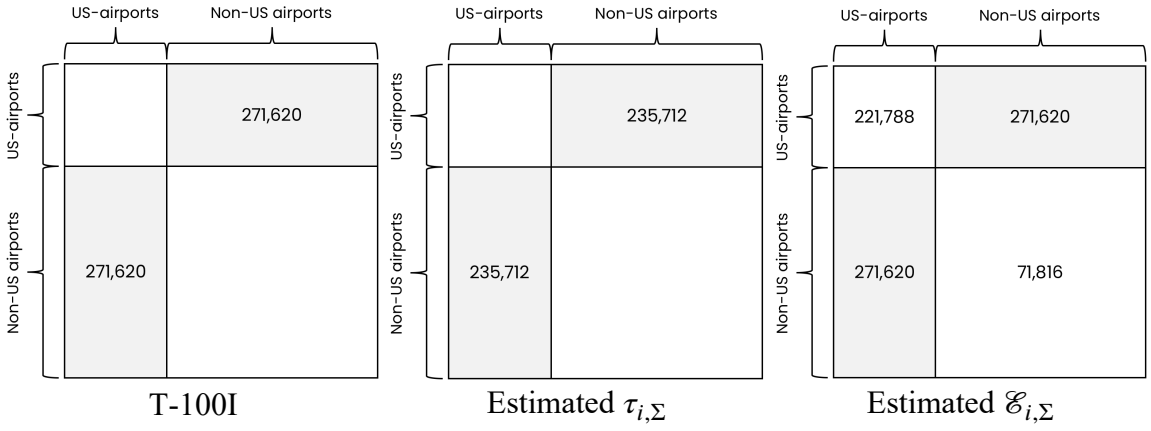


Figure 60 – Demand estimation result.

Finally, it is assumed that this demand also follows its smooth extrapolation. The total demand grows up to 1.75 times in 2050. Figure 61 visualizes the considered airports.

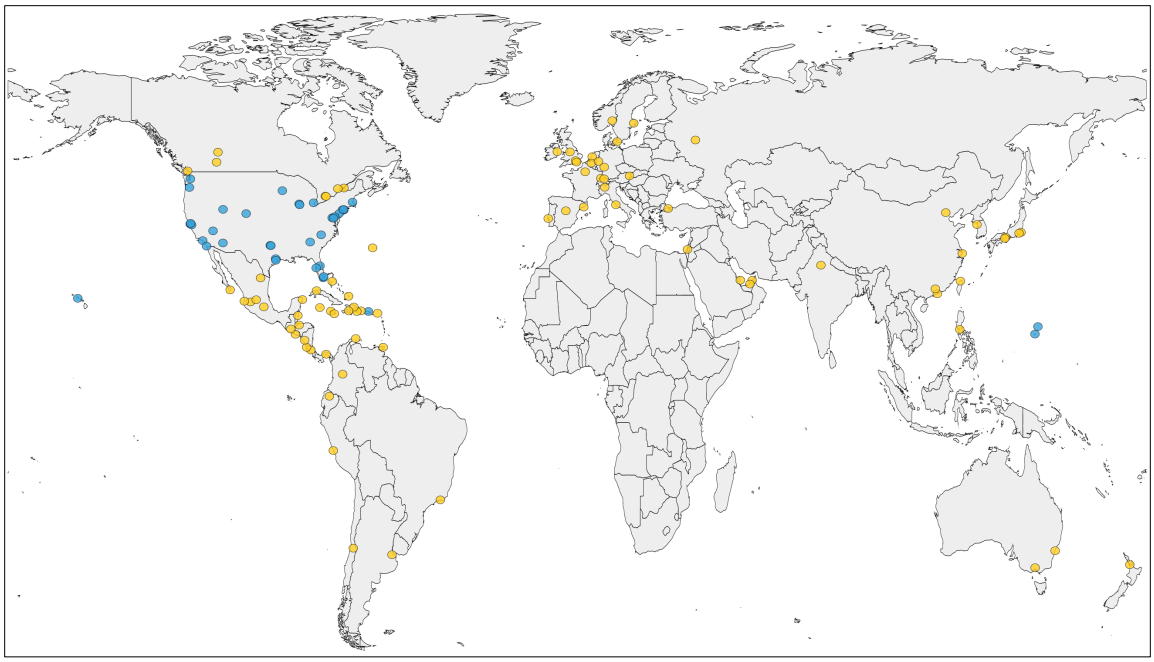


Figure 61 – Considered airports for international SST network simulation.

6.2.2 Considered Supersonic Aircraft Type

Table 31 summarizes the information about considered notional SST. [151]

Table 31 – Considered notional supersonic transport vehicle.

Capacity	V_{sup} (M)	Range (mi)
55	2.2 (1,452 mph)	4,500

6.3 Simulation

The finalized SST network evolution algorithm is organized as the following procedures. Note that the SST network evolution starts in 2025.

1. Perform ATN evolution of the core network
2. Identify all viable O-D segment

3. For all identified O-D trips,
    a. Identify all shortest path for all O-D SST travels
    b. Identify all paths where SST can operate: $range_{SST} > path_{i,j}$
    c. Calculate all O-D travel time savings via SST: $\Delta ft_{i,j}^{SST}$
4. Construct an SST network

As the developed architecture model was built for the domestic U.S. ATN, evolving the international core network has several following principles:

- Only the demands between non-U.S. airports and U.S. airports are considered in the T-100I so that all international flights which are not from or to U.S. airports cannot be modeled.
- The overall flight distances of international air travel are relatively much longer than those of domestic ones. Minimizing the flight duration is the top priority; the P2P-driven network is preferred in the network evolution.
- The number of flight routes should be one or two at maximum to augment the realism: γ is set as a very high value.

The international travel in validation considers the sum of total volume among the selected airports are considered such that $\mathscr{E}_{\Sigma} = \sum_{i<38, j\geq 38} \mathscr{E}_{i,j} + \sum_{i\geq 38, j<38} \mathscr{E}_{i,j}$, which is the sum of the T-100I, as represented in Figure 60.

6.3.1 Evolution of Components

The demand for the international SST network is also identified by the minimum threshold, 20 passengers per segment. This is less than that in the domestic SST network evolution

considering the volume difference of the T-100D and T-100I. Table 32 summarizes the evolution of the demand of the core network $\left(\sum_j \tau_{i,j}\right)$, international SST network.

Table 32 – Evolution of international ATN demand.

Year	τ_{Σ}^{Ref}	τ_{Σ}^{SST} (%)	$\left\|A^{SST}\right\|$
2017	471,430	7,619 (1.65)	-
2025	541,451	10,319 (1.91)	69
2030	604,541	11,972 (1.98)	73
2040	717,406	15,614 (2.18)	79
2050	839,912	20,196 (2.40)	100

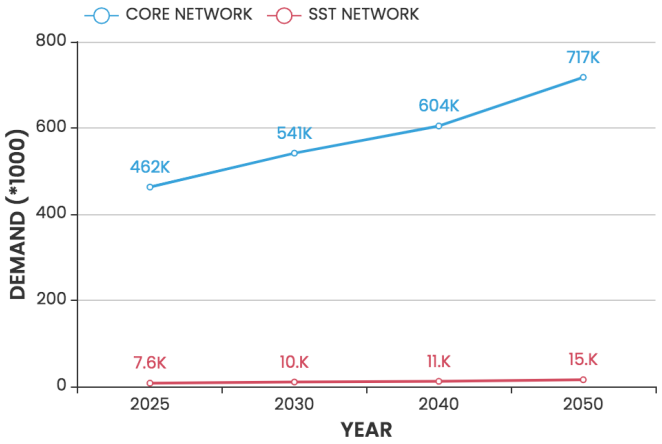
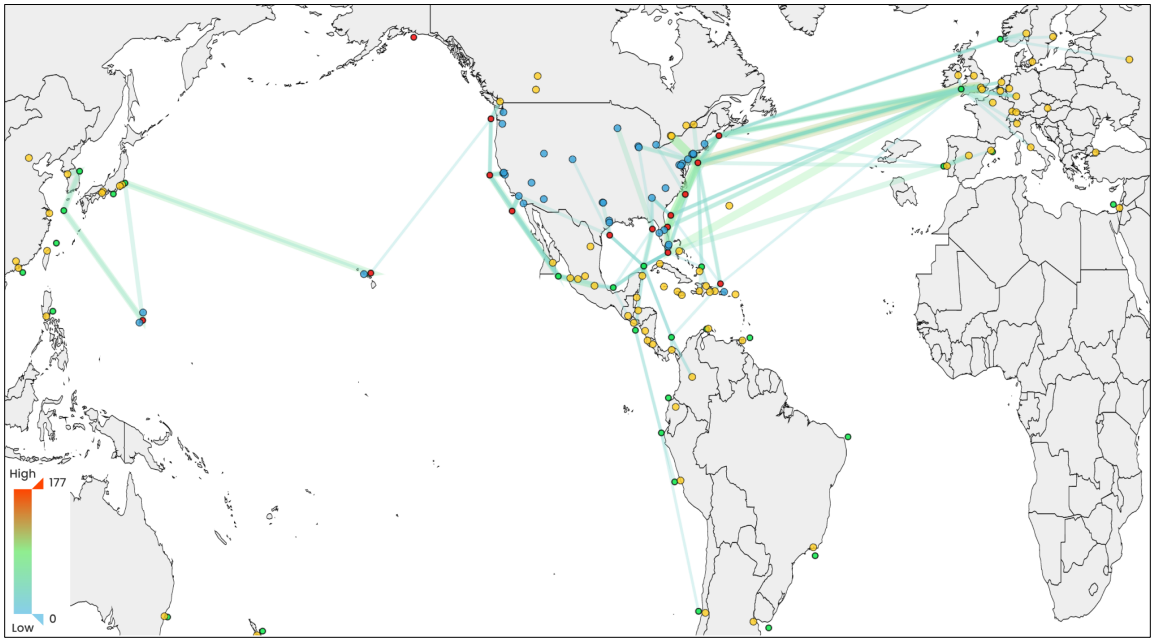


Figure 62 – Evolution of international ATN demand.

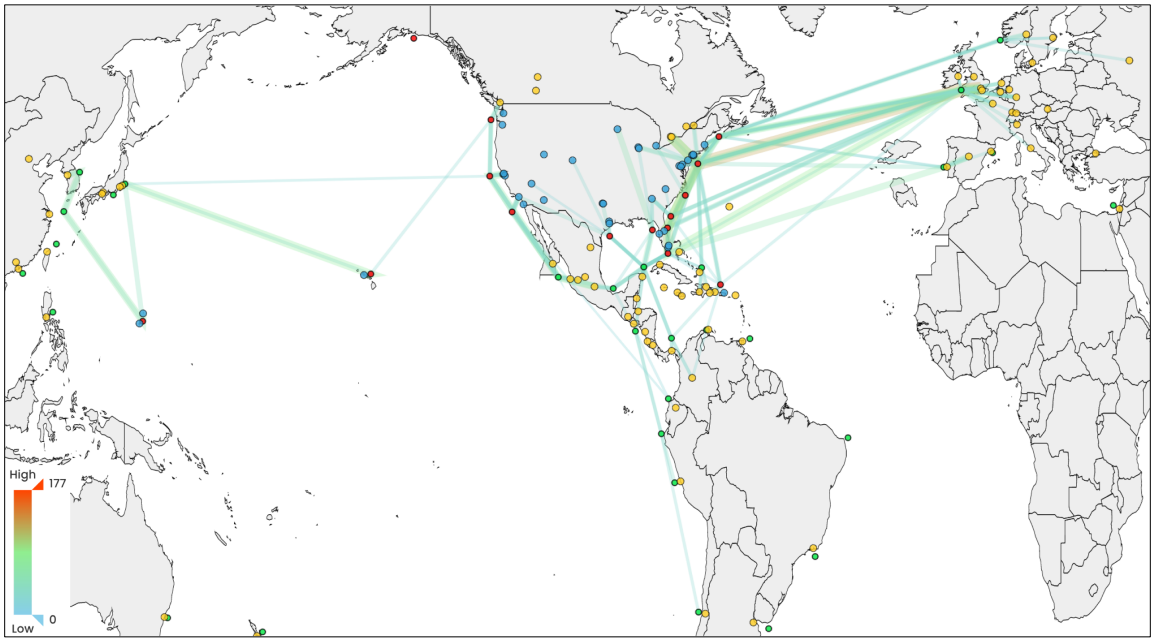
According to the demand forecasting, the total portion of the demand evolution is around 2% for all years. However, as the demand of the core network evolves, more airports become viable for the international SST network (see $\left|A^{SST}\right|$ in Table 32). This evolution is mainly dependent on the viability constraint, implying that it can be a new design parameter for controlling the future evolution of the need for SST travels throughout the world.

6.3.2 Evolution of International SST Network

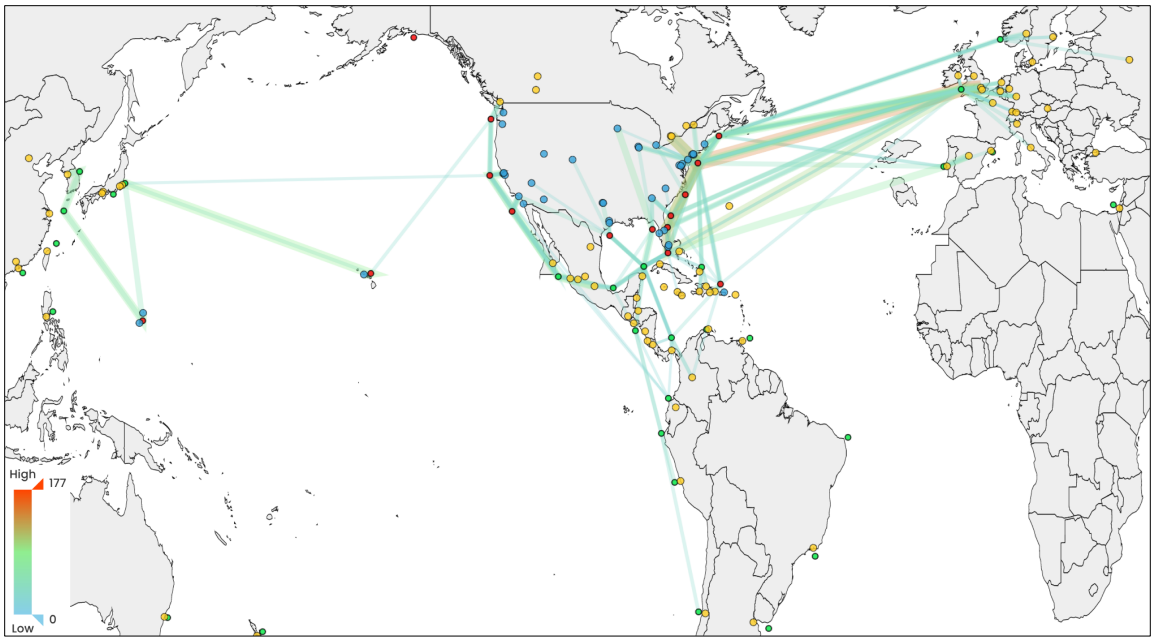
Figure 63 shows the SST network topologies, each of which was constructed for all selected evolution years.



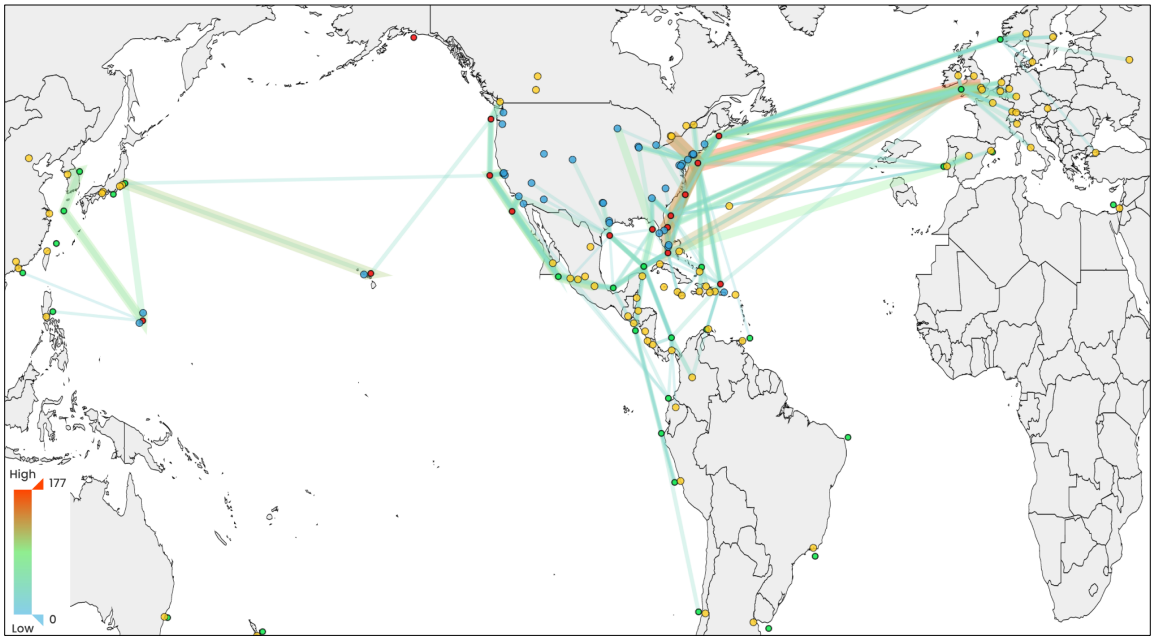
2025 $\left(\max\left(\mathscr{E}_{i,j}\right) = 110, \mathscr{E}_{\Sigma}^{SST} = 5{,}180, \left|A^{SST}\right| = 59\right)$



2030 $\left(\max\left(\mathscr{E}_{i,j}\right) = 128, \mathscr{E}_{\Sigma}^{SST} = 6{,}821, \left|A^{SST}\right| = 63\right)$



2040 $\left(\max\left(\mathscr{E}_{i,j}\right) = 147, \mathscr{E}_{\Sigma}^{SST} = 8{,}567, \left|A^{SST}\right| = 67\right)$



2050 $\left(\max\left(\mathscr{E}_{i,j}\right) = 177, \mathscr{E}_{\Sigma}^{SST} = 12{,}114, \left|A^{SST}\right| = 80\right)$

Figure 63 – Evolution of international SST network.

The topological analysis shows the growth of the international SST networks constructed by each SST vehicle allows envisioning the future ATN in the world. Table 33 summarizes the evolution of the international SST network.

Table 33 – Summary of international SST networks.

Year	$\mathscr{C}_{\Sigma}^{SST}$ (pax)	$\max(\mathscr{C}_{i,j})$ (pax)	Avg $\Delta ft_{i:j}^{SST}$ (hr)	$\mathscr{C}_{i,j}^{SST} \times \Delta ft_{i:j}^{SST}$ (hr)	Operations
2025	5,180	110	-2.470	-13,240	120
2030	6,821	128	-2.464	-17,319	72
2040	8,567	147	-2.469	-21,675	92
2050	12,114	177	-2.506	-30,750	128

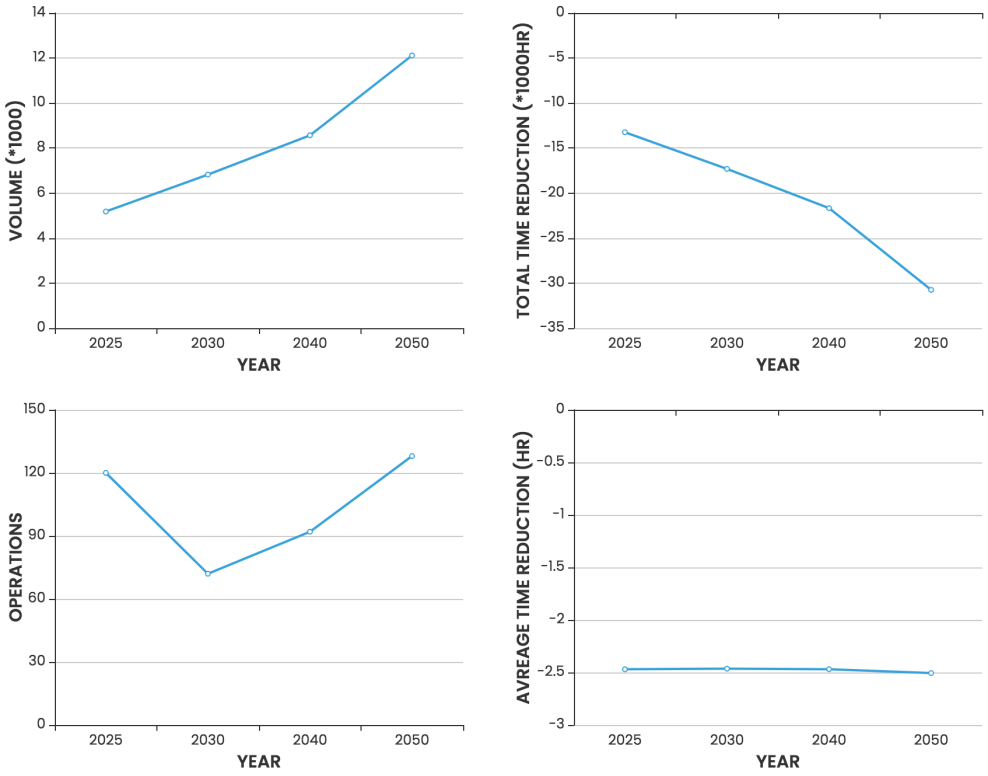


Figure 64 – Charts of international SST network construction summary.

The capacity issue is very crucial for accommodating the need for supersonic air travels. In this simulation, the number of viable aircraft fleet is assumed unlimited, but the reality will be quite different from the simulation result. If the expected delivery as the function of time is considered, then it would be another viability constraint, thereby recalling the importance of identifying the optimum routing for minimizing the disutility and maximize the profit.

The amount of travel time savings shows that SST has the potential for serving international air travel. It is much faster than any subsonic aircraft types so that the longer the distance becomes, the higher the level of passenger satisfaction is expected to be enhanced. Therefore, international flights are more attractive than domestic flights to fly supersonic due to the clear distinction of travel time savings. However, the total amount of demand for international flights is much smaller than that of the domestic markets. Therefore, the growth of induced demand for international markets seems to be one of the critical factors for the international SST network to evolve in the future.

6.4 Chapter Summary

So far, the future has been forecasted by using just no more than a linear extrapolation of enplanement and daily operations. [19] However, this approach can seldom allow researchers to forecast the future of the ATN in a multi-dimensional manner since the ATN is a huge system-of-systems with numerous constituents encompassing people, technology, economy, and various stakeholders.

Since all research questions were successfully answered, this chapter focused on envisioning the future deployment of the ATN evolution. To that end, a highly expected scenario, the emergence of the supersonic premium air transportation network, was chosen. A general notional SST vehicle under research and development was employed for the simulation.

The first procedure was to establish a set of design parameters associated with the SST technology. As a result, several design architectures were made: demand viability, range viability, buffer zone offset. Then, the evolutionary components were relevantly

modified. The core demand was extrapolated smoothly, airports were set invariant to time (i.e., no birth/demise of airports), and the selected supersonic transport was newly introduced for the SST network evolution.

Then, the design architecture was modified to embrace the evolution of the SST network. First, the devised viability factors were implemented into the main algorithm. Second, the network construction policy for supersonic route choice was built by heavily relying on the simplified shortest-path-finding approach, which explicitly engaged a set of waypoints for supersonic cruising (i.e., like docking points between subsonic and supersonic) to adequately incorporate the subsonic and supersonic cruises altogether. The simulated SST network was interpreted by metrics such as volume.

There do exist many limitations and over-simplifications in the research performed in this chapter. Above all, the lack of financial information of the SST would matter when considering the fleet assignment of multiple SST options. Despite this limitation, what has been accomplished in this chapter successfully simulated the evolution of the core network and the SST network internationally. The exploratory and interpretative capability of the architecture model has been tested through forecasting the future.

CHAPTER 7. CONCLUSION

CHAPTER 7 finalizes the dissertation by providing the summary, review, contributions, and recommendations for future work.

7.1 Research Summary & Review

This thesis dissertation has risen to the challenge of representing the complex ATN with its most parsimonious terms. To that end, the design architectures have been established in two interlocking perspectives: components and the dynamics. For the components (airports, demand, and aircraft types), the multi-dimensional comprehensive evolutionary information has been successfully established. As for the rules, a multi-tier network evolution approach has been developed, decomposing the entire ATN into its primary and secondary. The primary network construction has been driven by the proposed multi-objective probabilistic demand distribution algorithm mainly for the strong H&S structure, whereas the gravity-inspired modeling approach has governed the secondary network construction. Then, the full-scale network has accurately emerged while the evolutionary scheme was embracing these critical components and the rules albeit their forms being simplistic.

In this thesis, three research questions have been addressed from the identified research gaps. The research objective is revisited as follows for review:

RESEARCH OBJECTIVE: To develop an architecture model that can mathematically represent the evolution of the U.S. air transportation network led by the components and rules with realism.

The first research question and hypothesis focused on figuring out a better methodology to formulate the evolutionary information of components with improved realism compared to established ATN topology design models. The second research question and hypothesis strive to mathematically represent the rules of ATN evolution based on the second research gap and second observation: an aggregated single airline that deploys multiple network construction policies. The third research question and research hypothesis provided a broad approach to validate and verify the developed architecture model. The actual simulation from 1917 to 2018 resulted in a U.S. ATN topology that shows various essential aspects of the function of the real-world reference ATN as demonstrated by three research experiments.

Finally, a case study has been performed: forecasting the future ATN disruption caused by the technological revolution of civil supersonic transports. It provided an opportunity to experience the exploratory and interpretative capability of the architecture model, which shed light on performing future researches with better realism.

7.2 Research Contributions

This section provides the key contributions of this thesis:

- Development of an architecture model with a small number of design parameters rather than thousands of modeling parameters:

- This research succeeded in representing the evolution of the ATN by using its most parsimonious terms. To that end, a small number of design parameters were used to model the entire evolution of the ATN. Some were for prescribing the primary network structure while others were for determining specific constraints & thresholds.

- Introduction to the multi-tiered network evolution approach:
  - This research developed a multi-tiered network evolution approach that harnesses the synergism of decomposing the ATN into its multiple sub-tiers, each of which was governed by appropriate network construction algorithms. An elaborated multi-objective probabilistic demand distribution model for constructing the H&S structure was implemented to evolve the primary network tier, whereas a gravity-inspired approach for creating the multi-stop minor travels was devised to evolve the secondary network tier.
- Simulation-based ATN design & analysis framework:
  - This research established a simulation-based framework that can simulate different network evolutions based on various scenarios for the past, the current, and the future and allowing in-depth analysis of the created network topology.
- Evolutionary information of airports:
  - This research established the comprehensive historical information of all 438 U.S. airports from 1917 (DAL) to 2011 (SGU) for augmented realism.
- Evolutionary information of demand/volume:

- This research accomplished creating full origin-destination demand pairs from public itineraries by taking up to two-stop ones, thereby covering 99.9% of total network volume.

- Evolutionary information of aircraft types:
  - This research formulated the comprehensive information of the technological advancement of all aircraft fleets for augmented realism.
- Consideration of airport capacity constraint:
  - This research implemented airport capacity constraints in modeling approaches to adopt network disruption due to airport capacity. Due to the lack of timetable, daily operations were used. It is recommended that the hourly operation capacity can be tackled in future research.
- Handing of complex demand history (increase/decrease):
  - This research realized the capability to handle the historical fluctuation of the total amount of trip demand, including both increasing and decreasing.
- Demonstration of abstracting & grouping aircraft types:
  - This research accomplished in abstracting heterogeneous aircraft datasets using unsupervised machine learning algorithm.
- Explorable simulation of future supersonic network evolution:
  - This research successfully performed forecasting the future emergence of the supersonic international civil aviation network market by applying a set of assumptions and rationalizations. The elaborated design architecture was able to test the incubation and evolution of the supersonic market as well as the disruption of the established subsonic ATN.

7.3 Recommendations for Future Work

The proposed architecture model was accomplished with its potentials for future research. Including the components and dynamics, future works can be conducted in diverse directions.

- Different evolution path scenarios of components: One example is to forecast the possible future states of the ATN in response to hypothesized trends in demand, technology, policy, which could subsequently assist stochastic planning for operations and logistics. Moreover, possibly all what-if scenarios can be tested: different deployment of the evolution path, modeling & simulation on the temporary disruption of the ATN by different phenomenal events, and evaluating the structural change against the varying regulations, to name a few.
- Implementation of timetable and operations in a more granular fashion: The currently developed model considered the airport capacity, but the detailed timetable was beyond the scope of this thesis. As such, the hourly capacity constraint of airports has not been adequately considered. Instead, the daily number of operations was used as the target value of the constraint. Thus, in future research, elaborating the design architecture to embrace the operation scheduling of airports could be tackled. Once this is accomplished, then the appropriate airport capacity, the hourly maximum allowable number of operations could also be tackled.
- High-fidelity aircraft mission analysis module: The current architecture model focused on the demonstration of representing the complex ATN. The scope of

this thesis was limited. Thus, a simple short-haul model was used in estimating the aircraft operation cost. However, enhancing the fidelity of the aircraft mission analysis module is quite essential to improve the model's veracity as the airline cost is one of the objectives airlines consider in constructing the primary network tier. In future research, the current method could be more elaborated, or external high-fidelity tools such as the FAA's AEDT could be considered for augmented realism.

- Routing: This thesis did not consider aircraft routing; only the segmental round flights were considered. However, aircraft routing is also vivid evidence of the fact that airlines strive to minimize their operation costs. Thus, in future research, considering the routing operation could be much conducive to augment the realism of the architecture model.
- Modeling competitions of multiple airlines: This thesis simplified the chaotic game-theoretic competitions of different airlines into one single aggregated super-agent airline by a set of rationalizations. The reinforcing H&S structure has been manifested mainly by the airlines' struggles for survival. Hence, tackling this limitation by gradually increasing the number of abstracted airlines could be an important topic for future research.
- Integration of domestic and international ATN: The main scope of this thesis was the domestic network, represented by the DB1B and T-100D. In domestic flights, many international passengers are also included. That said, establishing the full O-D demand matrix information only from the DB1B could have an error to a certain extent. Knowing that no publicly available data source

provides the international itineraries, further research could also integrate the T-100I or other equivalent data sources accordingly. Then, it is expected that the small portion of the blended international passengers from/to the U.S domestic network will be able to be considered altogether, thereby allowing the model to simulate the evolution of the domestic and international ATN simultaneously.

Elaboration of the multi-tier network evolution approach: This thesis successfully proposed a multi-tier network evolution approach to represent the ATN. To enhance the adequacy of each sub-network construction method, the network was decomposed into the primary tier and secondary tier. The limitation is that the discrete separation of airports and demand into two different tiers was not natural. Even though there exist top major airports and regional airports that accommodate the demand in different dynamics, their clear distinction is still elusive. In this perspective, future research can perform not only elaborating on each algorithm but also increasing the number of network tiers with those advanced network construction approaches which will be developed in the future.

APPENDIX A. IMPORTANT NETWORK TERMS

A.1 General Graph Theory

Graph theory [152,153,154,155] has established the academic foundation of understanding a variety of complex entities in nature. In theory, any networks can be construed as a graph of which fundamental constituents are nodes (vertices, or points) and segments (edges, links, arcs, or lines), so a graph G is denoted as $G = (N, L)$, where $N \equiv \{n_1, n_2, n_3, \cdots, n_{|N|}\}$ and $L \equiv \{l_1, l_2, l_3, \cdots, l_{|L|}\}$ are the set of nodes (n) and the set of segments (l), respectively, and $|N| \geq 1$ is the size of the graph (e.g., the cardinality of N). Depending on the directionality of the graph, the total possible number of segments can be different as represented in the following equation:

$$\max_{1 \leq K = |N|} |L| = \begin{cases} K(K-1) & \text{if directed } G \\ K(K-1)/2 & \text{if indirected } G \end{cases} \tag{A.1}$$

This case corresponds to the case when all nodes are connected. Later, this equation is importantly employed in calculating any metrics associating with the clustering coefficient, which will be dealt with in the subsequent sections.

A.1.1 Adjacency

An adjacency is a mathematical expression for a network to identify the connectivity between two nodes. The matrix representation of the adjacency for a network of N nodes ($\mathscr{A}$) is denoted as follows:

$$\mathcal{A} = \left(a_{i,j}\right) \in \mathbb{R}^{N \times N}, \tag{A.2}$$

where $a_{i,j}$ is the element of the adjacency matrix $\mathcal{A}$ at i-th row and j-th column indicating whether node i (n_i) and node j $\left(n_j\right)$ are linked or not; if they are connected (adjacent), $a_{i,j} = 1$ and 0, otherwise. For example, if a segment between node 3 and 7 exist, $a_{3,7} = 1$. The adjacency is the foundational metric for other advanced metrics based upon it as it is the simplest one.

A weighted adjacency is distinguished from the basic adjacency in that it instead involves scalar weight that signifies some distinguishing traits such as distance and the number of operations. In mathematical formulation, the weighted adjacency matrix of a network ($\mathcal{A}^w$) is represented as follows with the superscript w:

$$\mathcal{A}^w = \left(a_{i,j}^w\right) \in \mathbb{R}^{N \times N} \tag{A.3}$$

A.1.2 Degree

A degree of node i (k_i) is the number of segments incident with the node, and is defined in terms of the adjacency matrix $\mathcal{A}$ as the following equation:

$$k_i = \sum_{j=1, j \neq i}^{N} a_{i,j} \tag{A.4}$$

As identified in the ATN analysis literature, the distribution of degrees is the top-level metric that allows researchers to characterize the structure of a network. For example, a

random network by the Erdős-Rényi model, [29] a scale-free network by the Barabási-Albert model, [1] and a small-world network by the Watts-Strogatz model [28] use the node degree distribution to distinguish from each other.

A.1.3 Strength

A strength of node *i* ($\mathcal{s}_i$) is the sum of the weighted adjacency of all other nodes associated with node *i* as denoted as below:

$$\mathcal{s}_i = \sum_{j=1, j\neq i}^{N} \mathcal{a}_{i,j}^{w} \tag{A.5}$$

It is the natural generalization of node *i* with degree $\mathcal{k}_i$ in various literature [156,157] by applying different information on the segments. In the ATN, the most representative weight of enplanements. Most metrics in this dissertation will be in the form of node strength containing different weights since the ATN is a complex network where airports topologically have a great deal of importance, emphasizing the importance of the aircraft as the network enabler.

A.1.4 Clustering Coefficient

The clustering coefficient is a metric that characterizes the acquaintance of a network by evaluating how two different nodes with a common neighbor are likely to be also neighbors to each other. [152,28] In other words, it is the number of triangles centered on that node divided by the number of triples centered on that node. The clustering coefficient of node *i* ($\mathcal{c}_i$) is calculated by the ratio of the actual number of segments among the neighbors of

node i (e_i) to the maximum possible number of segments among them, as shown in the following equation:

$$c_i = \frac{e_i}{k_i(k_i - 1)/2} = \frac{2}{k_i(k_i - 1)} \sum_{j=1, j\neq i}^{N} \sum_{m=1, m\neq i,j}^{N} a_{i,j}\, a_{j,m} a_{m,i} \tag{A.6}$$

As noticed, the word 'neighbor' implies no direction so that the number 2 in the numerator is applied. A clustering coefficient can be $0 \leq c_i \leq 1$. If $c_i = 0$, no two neighbors of n_i are neighbors to each other, else if $c_i = 1$, all neighbors of n_i are neighbors to every other neighbor of n_i. As for the weighted clustering coefficient $\left(c_i^w\right)$, the mathematical expression is represented as below [44]:

$$c_i^w = \frac{1}{s_i(k_i - 1)} \sum_{j=1, j\neq i}^{N} \sum_{m=1, m\neq i,j}^{N} \frac{a_{i,j}^w + a_{i,m}^w}{2} a_{i,j}\, a_{j,m} a_{m,i} \tag{A.7}$$

A.2 Complex Network

A.2.1 Centrality

The notion of centrality which originated from the social network analysis research [158] is another important research field in graph theory since it is a matter of identifying the most important nodes in a graph. [152,153] Depending on the types of networks, the most important nodes can have different meanings. In an ATN, they are mostly the top major airports, in a social network, they are the most influential person(s), and in a disease spreading network, they would be the ones who spread the disease fastest. The main

difference between the node influence metrics and centrality metrics is that the former is to measure the individual node influence in the network, whereas the latter seeks to quantify the relative importance among the nodes as if it calculated the centroid of a network.

Since this thesis is devoted to tackling the ATN which shows a firm H&S structure, tackling the various aspects of the function of the ATN via various centrality metrics will be thoroughly conducted with those introduced. The centrality metrics involve the node influence metrics in their being calculations so that in general, calculating a centrality requires more expensive than calculating a node influence metric as the centrality needs to explore and compare all nodes in a network. However, network scientists want to focus on the most important nodes or segments in the complex network structure. Therefore, centrality can provide the solution for them so that the relatively higher computational cost is worthy of conducting.

It is highly conjectured that centrality metrics, in most cases, will identify the top major airports as the most important airports. Moreover, their historical variations in the ATN evolution will also be importantly investigated to prove the research hypotheses.

A.2.1.1 Shortest Path

The shortest path can be based on either just the number of segments (binary adjacency: $a_{i,j}$) or the weight of segments in the path (weighted adjacency: $a_{i,j}^{w}$). If betweenness uses the former (assuming that all nodes can traverse to all others), then there can be a multitude of the shortest paths between n_s and n_t such that $\sigma_{s,t} \geq 1$, while if the latter is used, the shortest path is often unique due to the different weights of segments in the path.

Concerning this, this thesis adopts what has been proposed by Opsahl et al. [159] to calculate the weighted shortest path in the ATN. As such, the total route weight of $r_{v_1:v_l}$ $\left(\mathscr{R}_{v_1:v_l}\right)$ is expressed as the following equation:

$$\mathscr{R}_{v_1:v_l} = \mathscr{R}_{v_1,v_2,\cdots,v_{l-1},v_l} = \begin{cases} \sum_{i=1}^{l-1} a^w_{v_i,v_{i+1}} & \text{if } a^w_{v_i,v_{i+1}} \text{ is weight} \\ \sum_{i=1}^{l-1} \frac{1}{a^w_{v_i,v_{i+1}}} & \text{if } a^w_{v_i,v_{i+1}} \text{ is cost} \end{cases} \tag{A.8}$$

If the weighted adjacency $\left(a^w_{v_i,v_{i+1}}\right)$ is actual weight or importance (i.e., enplanement, discount), then $\mathscr{R}_{v_1:v_l}$ is the sum of the weight of all segments in $r_{v_1:v_l}$. On the contrary, if $a^w_{v_i,v_{i+1}}$ is a cost which implies the smaller one is preferred (i.e., distance, operation cost, flight time), then the sum of the inverses is considered. Eventually, the weighted shortest path of $r_{v_1:v_l}$ $\left(\mathscr{R}^w_{v_1:v_l}\right)$ is identified by solving the optimization problem:

$$r^w_{v_1:v_l} = \underset{r_{v_1:v_l}}{\operatorname{argmin}} \mathscr{R}_{v_1:v_l}, \tag{A.9}$$

where the corresponding weighted shortest path of $r^w_{v_1:v_l}$ is denoted as $\mathscr{R}^w_{v_1:v_l} = \min \mathscr{R}_{v_1:v_l}$. There are numerous heuristic algorithms to find the shortest path. In this thesis, the Dijkstra's algorithm is used. [144]

A.2.1.2 Betweenness Centrality

Betweenness is a centrality metric that matters how many times a node or a segment lies on all shortest paths of all node pairs in a network. It is one of the standard centrality metrics

to quantify the importance of network constituents and has been firstly proposed by some insightful papers. [160,161,162] According to Linton's research, [160] nodes with a high likelihood to lie on a randomly sampled shortest path between two individual nodes also have a high betweenness. From a physical perspective, betweenness centrality is that it quantifies the controllability for a node to exert over other nodes in a network. The interactions of others. [163] The betweenness of node i $\left(\mathscr{C}_{B,i}\right)$ is calculated by the following equation:

$$\mathscr{C}_{B,i} = \sum_{s=1,s\neq i}^{N} \sum_{t=1,t\neq i,s}^{N} \frac{\sigma_{s,t}(i)}{\sigma_{s,t}}, \tag{A.10}$$

where $\sigma_{s,t}(i)$ represents the number of occurrences where n_i lies on $r_{s:t}^{w}$ and $\sigma_{s,t}$ is the number of $r_{s:t}^{w}$. In the same way, the betweenness centrality of $s_{i,j}$ $\left(\mathscr{C}_{B,i,j}\right)$ is calculated as follows:

$$\mathscr{C}_{B,i,j} = \sum_{s=1,s\neq i,j}^{N} \sum_{t=1,t\neq i,j,s}^{N} \frac{\sigma_{s,t}(i,j)}{\sigma_{s,t}}, \tag{A.11}$$

where $\sigma_{s,t}(i,j)$ represents the number of occurrences when $s_{i,j} \in r_{s,t}^{w}$.

A.2.1.3 Closeness Centrality

Closeness centrality was firstly proposed by Alex Bavelas [164] and has been elaborated in various researches. [165] It is a matter of the swiftness of flow in a network. In an information network, it evaluates how fast information is conveyed between two nodes. In

air transportation, the closeness centrality of airport *i* assesses the reciprocal of the average cost/weight of the weighted shortest paths from it to all other airports. For example, the transportation is very dependent on the distance so that if the weighted adjacency was simply the distance, an airport with a high closeness would be, intuitively, at a spot in the vicinity of the geographical centroid of all airports in the ATN. Thus, the more central a node is, the closer it is to all other nodes. Differently from betweenness centrality, it is not evaluated on a segment. Finally, the closeness centrality of node *i* $\left(\mathscr{C}_{C,i}\right)$ is calculated by the following equation [159]:

$$\mathscr{C}_{C,i} = \frac{1}{\langle \mathscr{R}^{w}_{i:j} \rangle} = \frac{N}{\sum_{j=1, j\neq i}^{N} \mathscr{R}^{w}_{i:j}} \tag{A.12}$$

As described, closeness centrality is another metric to identify the relative importance of airports. As the top major airports have a lot of direct routes, they will have high values of closeness. In a perfect P2P network, the $\mathscr{C}_C$ for all airports is 1 (using binary adjacency), indicating that the average shortest path length from an origin to all destinations is just one segment to fly: direct only.

A.2.2 Scale-Free Property

Scale-free property checks the cumulative degree distribution, the most common binary adjacency throughout airports. In many cases, the distribution is regressed via different mathematical functions depending on the networks' topological characteristics. In case of the strong H&S showing the scale-free property, it has been known that the degree distribution exhibits the heavy-tails (i.e., the power-law) by the research of Barabasi and Albert. [1]

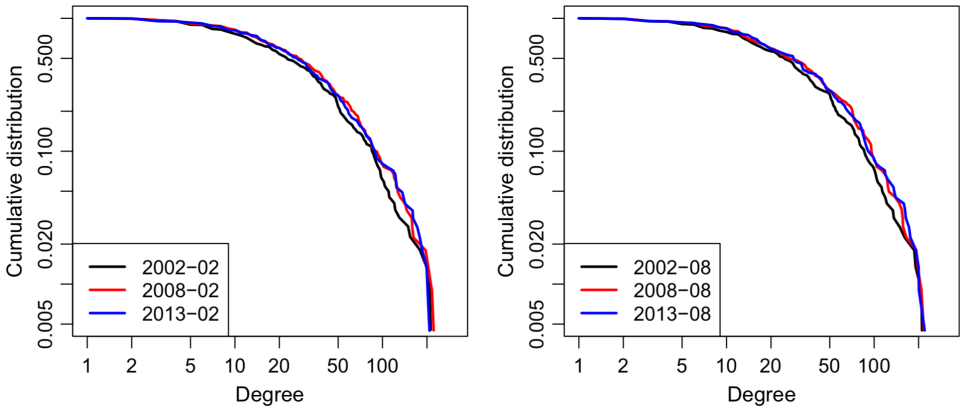


Figure 65 – Example visualization of scale-free property. [32]

A.2.3 Small-World Property

Small-world property is characterized by comparing the two-dimensional population of the average shortest path length $\left(\langle \mathscr{R}_{v_1:v_l} \rangle\right)$ and clustering coefficient (c_i) of the target network with the random network having the same number of segments and nodes to the target network. The node-wise average shortest path length is the ordinate, and the clustering coefficient is the abscissa. Both of binary and weighted adjacency can calculate both metrics. The basic definition is established by the seminal research of Erdos and Renyi, [29] and the comparison with a random network can vividly confirm small-world property.

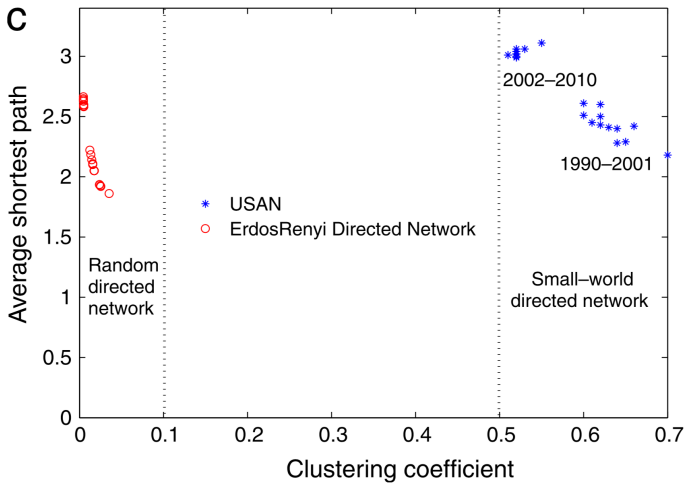


Figure 66 – Example visualization of small-world property. [57]

A.3 Transportation Network

A.3.1 Triads Census

A triad in a network is a unique 3-node substructure in a network. [166,167] If the number of occurrences of each triad in a network is called triad census. Faust [168] proves that there are 16 distinct types of possible triads, as shown in Figure 67. [32] The naming follows the standard 'MAN' notation firstly proposed one by Holland and Leinhardt. [169,170,171] Each letter of the MAN notation indicates the number of mutual (M), asymmetric (A), and null (N) dyads in each triad.

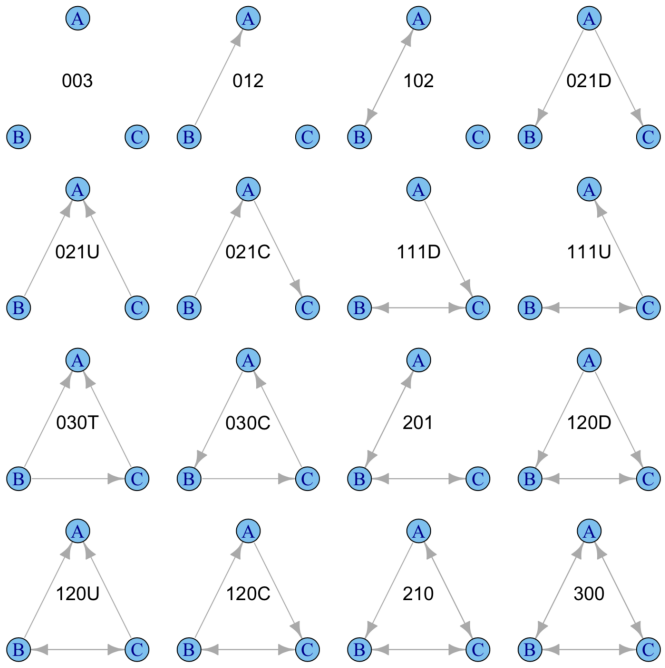


Figure 67 – Types of triads in a directed network.

For a symmetric network, the sixteen different types are curtailed into four types depending on the number of segments in a triad: 0-segment (003), 1-segment (012 and 102), 2-segment (021D, 021U, 021C, 111D, 111U, 201), and 3-segment (030T, 030C, 120D, 120U, 120C, 210, 300). Although the triad census is based on the binary adjacency, the metric associated with the segment can also be a weighted adjacency.

A.3.2 Gravity

In transportation network research, gravity is widely used to formulate the intrinsic transportation demand between two different cities. There are numerous empirical studies on the ATN demonstrating that the traffic between airports $\left(w_{i,j}\right)$ can be generally modeled by the following equation [172]:

$$w_{i,j} \sim x_{i,j} \left(\mathscr{A}_i^w \mathscr{A}_j^w \right)^\theta \tag{A.13}$$

Here, $w_{i,j}$ is the traffic between airport i and airport j, $x_{i,j}$ is a random real number, θ is a positive exponent, and $\mathscr{A}_i^w$ and $\mathscr{A}_j^w$ are the weighted adjacency of airport i and airport j, respectively. In more general form, it is represented in a form which is like the law of universal gravitation:

$$w_{i,j} = K \frac{\left(\mathscr{A}_i^w \mathscr{A}_j^w \right)^{\theta_1}}{{d_{i,j}}^{\theta_2}} \tag{A.14}$$

Here, θ_1 and θ_2 are positive tuning parameters and $d_{i,j}$ is the distance between the two airports, and K is a constant coefficient. Jung et al. [173] proved that this simple equation could be suitable for modeling a network typically constructed in a geographical space since parameters act as if they were particles. [174]

In this thesis, however, this equation is not used for architecting the ATN. Preferably, this equation is used to describe the gravity field in the CONUS by involving some important metrics including volume ($\mathscr{C}$), bandwidth ($\mathscr{B}$), and centrality. For all geographical coordinates encompassing the CONUS, the corresponding gravity vector denoted as G can be defined, forming up the entire gravity field manifested from the ATN. Then, by harnessing the vector representation, the relative gravity (i.e., the relative importance of all airports) could be identified. To assist readers in understanding this concept, an illustrative visualization is shown in Figure 68.

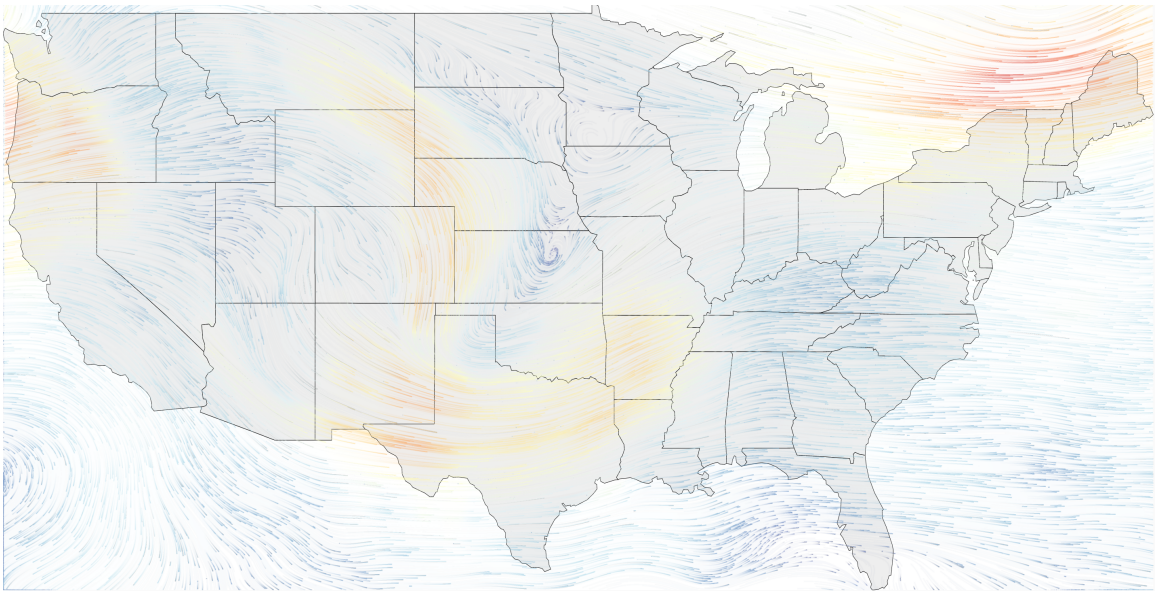


Figure 68 – Example vectorized visualization of gravitational flow in CONUS.

In Figure 68, the difference of gravity is represented by a set of gradient colors from blue to red, meaning low and high magnitude of the gravity, respectively. The gravity vector exerted by airport y on a coordinate point x $\left(\overrightarrow{\mathscr{G}_{x,y}}\right)$ is as follows:

$$\overrightarrow{\mathscr{G}_{x,y}} = K\frac{\left(\mathscr{A}_y^w\right)^{\theta_1}}{{d_{x,y}}^{\theta_2}} \times \frac{1}{dl_{x,y}}\begin{bmatrix} lon_y - lon_x \\ lat_y - lat_x \end{bmatrix}, \tag{A.15}$$

where x is a geographical point represented as its longitude (lon_x) and latitude (lon_x), $dl_{x,y}$ is the L2-norm of the vector $\begin{bmatrix} lon_y - lon_x & lat_y - lat_x \end{bmatrix}^\intercal$. lon_y and lat_y are the longitude and latitude of α_y, respectively. Note that there is only one weighted adjacency term $\left(\mathscr{A}_y^w\right)$ in the numerator of the equation. Dividing into two parts by the multiplication symbol in the middle ($\times$), the first part is the scalar magnitude of the gravity $\left(\left\|\overrightarrow{\mathscr{G}_{x,y}}\right\|\right)$ and the second part is the unit direction vector of gravity. Therefore, the net gravity vector exerting on a

coordinate point x in the CONUS associated with all airports is evaluated by taking the sum of all individual vectors which are incident upon the point:

$$\overrightarrow{\mathscr{G}_x} = \sum_{y=1}^{N} \overrightarrow{\mathscr{G}_{x,y}} = K \sum_{y=1}^{N} \left(\frac{\left(\mathscr{A}_y^w\right)^{\theta_1}}{{d_{x,y}}^{\theta_2}} \times \frac{1}{dl_{x,y}} \begin{bmatrix} lon_y - lon_x \\ lat_y - lat_x \end{bmatrix} \right) \tag{A.16}$$

As such, the final vector is determined as the outcome of the wrestling between the individual gravity exerted by each airport.

A.4 Air Transportation Network

A.4.1 Volume (Enplanements)

Enplanement or volume is the total number of passengers who board the aircraft to fly from α_i to α_j in a period of interest. This is the most simplistic yet important metric to measure the weight of the ATN. In this thesis, this will play a role as the main scalar metric to evaluate the model's veracity concerning the RQ3 and RH3. In mathematical form, the volume (of passengers) of segment $s_{i,j}$ and route $r_{i:j}$ are denoted as $\mathscr{E}_{i,j}$ and $\mathscr{E}_{i:j}$, respectively. Recalling that a route may consist of multiple segments, the total enplanement of an arbitrary route ($\mathscr{E}_{i:n}$) is calculated by abiding by the below equation:

$$\mathscr{E}_{i:n} = \mathscr{E}_{i,i+1,\cdots,n-1,n} = \sum_{j=i}^{n-1} \mathscr{E}_{j,j+1}, \tag{A.17}$$

where $n \geq 2$ and $r_{i:n} = r_{i,i+1,\cdots,n-1,n} = (i, i+1, \cdots, n-1, n)$ is assumed to be a sequence of $(n-i)$ segments. Conceivably, the total departure enplanement from α_i $\left(\mathscr{E}_{i,\Sigma}\right)$ and the total arrival enplanement to α_j $\left(\mathscr{E}_{\Sigma,j}\right)$ are denoted as $\mathscr{E}_{i,\Sigma} = \sum_{j=1,j\neq i}^{N} \mathscr{E}_{i,j}$ and $\mathscr{E}_{\Sigma,j} =$

$\sum_{i=1,i\neq j}^{N} \mathscr{E}_{i,j}$, respectively, where the sigma (Σ) in the subscript represents 'summing up all airports' in the ATN. Note that $\mathscr{E}_{i,\Sigma}$ and $\mathscr{E}_{\Sigma,j}$ correspond to the row-sum and the column-sum of the enplanement matrix, which is defined as $\mathscr{E} = \left(\mathscr{E}_{i,j}\right) \in \mathbb{R}^{N\times N}$. Notably, the departure enplanement and arrival enplanement will be dubbed production and attraction, respectively. Therefore, the i-th row of $\mathscr{E}$ is denoted as $\mathscr{E}_{i,-} = \left[\mathscr{E}_{i,1}\ \mathscr{E}_{i,2}\ \cdots\ \mathscr{E}_{i,N}\right]$ while the j-th column of $\mathscr{E}$ is denoted as $\mathscr{E}_{-,j} = \left[\mathscr{E}_{1,j}\ \mathscr{E}_{2,j}\ \cdots\ \mathscr{E}_{N,j}\right]^{\intercal}$, where the minus (-) symbol implies 'arbitrary airports'. The subscripts (e.g., ':', 'Σ', and '-') will be universal regardless of metrics; if a metric can be represented as a matrix form, then these subscripts convey the same notions.

A.4.2 Bandwidth

Bandwidth is developed to identify the relative importance of a segment. The importance of an airport heavily relies on the importance of the segments associated with it. In a segment, one of the typical phenomena of an H&S structure is the complex mixture of passengers whose origins and destinations are quite different. Therefore, the volume can intuitively assess how important a segment is.

However, thanks to the benefits of advanced aircraft and the substantial scale-free property in the ATN, the geographical distance between airports is not quite a significant matter in some top segment markets. In the US ATN, for instance, the majority of the top major airports reside in coastal lines of the CONUS (e.g., LAX, JFK, SEA, SFO) and the market between LAX and JFK airports is one of the most lucrative ones. According to the T-100D by the Bureau of Transportation Statistics (BTS) of 2018, [7] the daily

enplanement from LAX to SFO $(\mathscr{E}_{LAX,SFO})$ was 4,715 while that from LAX to JFK $(\mathscr{E}_{LAX,JFK})$ was 5,497. However, the distances are quite disparate: 337 miles and 2475 miles, respectively. That said, even though two different segments have the same volume, the one with a smaller flight distance would be certainly more important than the other.

Hence, considering that the ATN is a 'transportation' network that highly concerns the time and distance, there arises a need to develop a new metric to evaluate the importance of segments which can take into account the influence of distance. Inspired by this idea, a metric dubbed bandwidth is defined as the ratio of the volume to the distance of a segment. The bandwidth of $s_{i,j}$ $(\mathscr{B}_{i,j})$ is defined as the below equation:

$$\mathscr{B}_{i,j} = \frac{\mathscr{E}_{i,j}}{d_{i,j}}, \tag{A.18}$$

where $d_{i,j}$ is the geodesic distance of α_i and α_j. This equation is the number of passengers per unit flight distance so that it can be interpreted as a matter of how many passengers can fly at the same time for a specific market: a bandwidth of a segment. Recalling the two segments mentioned above, even though the enplanement says that the segment from LAX to JFK is 'slightly' more critical than the segment from LAX to SFO, the bandwidth refutes it by involving the distance as a new factor: $\mathscr{B}_{LAX,SFO} = 13.99$ and $\mathscr{B}_{LAX,JFK} = 2.22$. According to the bandwidth, the segment from LAX to SFO is approximately 5.3 times more important than the other. Finally, the bandwidth of α_i denoted as $\mathscr{B}_i$ is calculated by summing up all its associated bandwidths as expressed below:

$$\mathscr{B}_i = \sum_{j=1, j\neq i}^{N} \mathscr{B}_{i,j} = \sum_{j=1, j\neq i}^{N} \frac{\mathscr{E}_{i,j}}{d_{i,j}} \tag{A.19}$$

As such, the more connections does an airport have (high degree), the higher the bandwidth becomes. In the same way, the higher the enplanement does an airport have, the higher the bandwidth becomes.

APPENDIX B. AIRCRAFT TYPE MODULES

B.1 Updating Cost Data for Aircraft Type Groups

Because the abstracted ten aircraft types have different values of a performance specification, the fixed cost and variable cost must be modified accordingly. To that end, three columns of data are needed: hourly cost ($) and block time. The cost must be averaged based on the block time of each aircraft. The principle is simple; the more an aircraft type is utilized, the more reliable and important the cost becomes. The block time is also given by Schedule P-5.2. At last, the modified cost for each group of aircraft types is calculated in an array of mathematical procedures. To better convey the cost analysis process, Group 5 (G5, in short) is engaged as an example. Firstly, cost, block time, normalized block time (0 ~ 1 scale) from 2002 to 2017 are tabulated below. Note that the indices i and j represent the aircraft type (row) and year (column), respectively.

Table 34 – Example cost estimation of G5.

1. *Cost* (\$/hour: $c_{i,j}$)

Aircraft Type	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
McDonnell Douglas DC-8-71	36,017	8,821	10,676	12,908	12,428	13,933	16,235	10,935	13,307	16,136						
Airbus Industrie A300B/C/F-100/200	13,554	5,059	5,470	5,885	5,836	10,426	11,545	19,865								
McDonnell Douglas DC-10-30CF			15,268	8,975	7,238	8,506	13,372	9,233	10,217							
Airbus Industrie A310-200C/F	36,387	10,368	12,040	13,045	12,637	14,033	16,127	13,493	14,588	14,941	15,257	22,677	16,274	14,957	17,893	11,458
Boeing 757-300	7,719	3,710	3,777	4,346	4,698	5,316	6,665	5,370	5,122	6,384	6,933	8,293	7,206	5,114	4,856	5,048

2. *Block time* (hours: $bt_i = \sum_j bt_{i,j}$)

Aircraft Type	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	bt_i
McDonnell Douglas DC-8-71	6.31	20.05	19.49	17.44	19.35	16.35	12.21	6.99	5.97	3.98							128.1
Airbus Industrie A300B/C/F-100/200	2.26	10.22	8.43	8.2	7.61	4.83	4.42	1.9									47.9
McDonnell Douglas DC-10-30CF			0.22	6.07	8.39	9.99	9.77	9.88	5.48								49.8
Airbus Industrie A310-200C/F	14.02	53.57	52.66	55.89	65.27	65.73	56.37	39.24	38.82	41.37	34.02	25.41	21.81	16.73	7.44	11.68	600.0
Boeing 757-300	14.57	79.77	127.44	135.72	136.15	131.3	128.56	112.48	119.72	119.41	126.79	126.42	123.95	129.49	125.07	127.3	1,864.1

3. *Weighted cost* (×\$1,000/hour: $wc_i = \sum_j c_{i,j} bt_{i,j} / \sum_i bt_i$)

Aircraft Type	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	wc_i
McDonnell Douglas DC-8-71	227.3	176.9	208.1	225.1	240.5	227.8	198.2	76.4	79.4	64.2							13.454
Airbus Industrie A300B/C/F-100/200	30.6	51.7	46.1	48.3	44.4	50.4	51.0	37.7									7.525
McDonnell Douglas DC-10-30CF			3.4	54.5	60.7	85.0	130.6	91.2	56.0								9.667
Airbus Industrie A310-200C/F	510.1	555.4	634.0	729.1	824.8	922.4	909.1	529.4	566.3	618.1	519.0	576.2	354.9	250.2	133.1	133.8	14.610
Boeing 757-300	112.5	296.0	481.3	589.9	639.6	698.0	856.9	604.1	613.2	762.3	879.0	1,048.5	893.2	662.3	607.3	642.7	5.572

Table 35 – Summary of individual contribution of aircraft types in G5.

Aircraft type	Absolute block time (bt_i)	Weighted cost (wc_i)	Relative cost $\left((bt_i \times wc_i)/\sum_i bt_i\right)$
McDonnell Douglas DC-8-71	128.1	13,454	640.9
Airbus Industrie A300B/C/F-100/200	47.9	7,525	133.9
McDonnell Douglas DC-10-30CF	49.8	9,667	179.0
Airbus Industrie A310-200C/F	600.0	14,610	3258.8
Boeing 757-300	1864.1	5,572	3861.2

4. *Total cost*

$$Cost_5 = \sum_{i=1}^{5} wc_i = 640.9 + 640.9 + 179.0 + 3258.8 + 3861.2 = 8074.8 \text{ (\$/hour)}$$

Through this procedure, the finalized hourly cost of the aircraft types in G5 ($Cost_5$) is obtained as $8074.8 per hour. This is the global scheme to get the cost information for all aircraft type groups from G1 to G10.

The blue-colored cells in the table of procedure 2 show that among the five aircraft types in G5, Airbus Industrie A310-200C/F and Boeing 757-300 are the dominant ones. The total portion of their block time is approximately 0.87, being construed that these two aircraft types flew over the sky for 87% compared to other types in similar performance specifications. Therefore, $Cost_5$ is very similar to those two aircraft types. Otherwise, the outweighing cost of McDonnell Douglas DC-8-71 (See the grey cells in the table of procedure 1) would dominate but its relative block time is minimal (~0.05). Hence, its contribution to $Cost_4$ becomes minuscule.

B.2 Cost Estimation of G1 – Boeing 40A, Boeing 80A, and Boeing 247

The first group of aircraft types consists of Boeing 40A, Boeing 80A, Boeing 247, and Bell B-206A. Table 36 shows the basic cost data.

Table 36 – Cost data of G1 from Schedule P-5.2.

Aircraft Type	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Boeing 40A																
Boeing 80A																
Boeing 247																
Bell B-206A	6,489	1,405	1,855	1,058	705	2,036	2,158									

The first three aircraft types come with no cost information. Hence, the cost of G1 is set to be simply that of Bell B-206A. Focusing on the cost data of Bell B-206A, the grey-colored cell of 2002 is an outlier. Therefore, the final weighted cost of G1 ($Cost_1$) is calculated by

$Cost_1 = (1405 + 1855 + 1058 + 705 + 2036 + 2158)/6 = 1536$. Conceivably, the block time is assumed to be evenly distributed: 0.25 per aircraft type.

B.3 Cost Estimation of G3 – Boeing 707-120, Douglas DC-8, Boeing 720, and Convair 990

The second group of aircraft types comprises eight aircraft types: Boeing 707-120, Douglas DC-8, Boeing 720, Convair 990, Airbus Industrie A320-100/200, Airbus Industrie A321, Boeing 737-800, and Boeing 737-900. Table 37 shows the basic cost data.

Table 37 – Cost data of G3 from Schedule P-5.2.

Aircraft Type	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Boeing 707-120																
Douglas DC-8																
Boeing 720																
Convair 990																
Airbus Industrie A320-100/200	227	987	1,143	1,272	1,341	1,383	1,331	1,316	1,395	1,490	1,557	1,621	1,634	1,675	1,700	1,707
Airbus Industrie A321	25	115	116	136	143	126	124	150	183	190	239	280	430	621	873	1,104
Boeing 737-800	182	715	852	883	957	1,044	1,082	1,089	1,241	1,405	1,620	1,854	1,995	2,204	2,295	2,510
Boeing 737-900	19	76	88	89	85	85	113	157	183	194	238	353	522	624	613	630

Since the last four aircraft types have the entire years of cost data (filled by the blue color), it is obvious that the cost of each aircraft type is calculated by taking the average of each aircraft type, which \$4,070, \$4,075, \$4,120, and \$3,879, respectively. Integrated performance data is tabulated in Table 38.

Table 38 – Performance specification of aircraft types in G3.

Aircraft Type	Range	Speed	Capacity	Debut Year
Boeing 707-120	174	4,100	607	1959
Douglas DC-8	177	3,760	556	1959
Boeing 720	149	4,350	621	1960
Convair 990	149	3,595	557	1962
Airbus Industrie A320-100/200	195	3,798	528	1988
Airbus Industrie A321	190	3,685	516	1994
Boeing 737-800	160	3,378	523	1998
Boeing 737-900	177	3,393	541	2001

In order to estimate the cost of the first four aircraft types, the costs of the last four aircraft types are linearly regressed. The best-fitting equation for the linear regression is represented as the following equation:

$$Cost_{3,i} = 5000.1045973 - 10.94987444 \times Cap_i + 0.8380336102 \times Rng_i - 3.745702144 \times Spd_i, \quad \text{(B.1)}$$

where $Cost_{3,i}$, Cap_i, Rng_i, and Spd_i indicate the cost, capacity, range, and speed of i-th aircraft type in G3. Astonishingly, this simple linear equation perfectly regresses the costs of reference four aircraft types. Complete data and error are summarized in Table 39.

Table 39 – Summary of cost estimation using linear regression.

Aircraft Type	Range	Speed	Capacity	Cost (ref.)	Cost (est.)	Error
Boeing 707-120	174	4,100	607	-	4257.1	-
Douglas DC-8	177	3,760	556	-	4130.4	-
Boeing 720	149	4,350	621	-	4687.9	-
Convair 990	149	3,595	557	-	4294.9	-
Airbus Industrie A320-100/200	195	3,798	528	4070	4069.2	0.019%
Airbus Industrie A321	190	3,685	516	4075	4075.8	0.019%
Boeing 737-800	160	3,378	523	4120	4119.1	0.022%
Boeing 737-900	177	3,393	541	3879	3877.9	0.028%

As seen, the errors are quite ignorable. Therefore, without further contemplation, this equation is applied to estimate those which miss their cost data.

As to the block time, a modest approach is applied. Since the cost data of the first four aircraft types are estimated, it can barely say that they are as reliable as the cost data of the others. Hence, a small arbitrary block time ratio of 0.05 is imposed on them. This block time ratio is an essential factor in calculating the varying cost data for all groups of aircraft types through the ATN evolution deployed in the developed architecture model.

B.4 Aircraft Mission Analysis Module

The first requirement to perform to estimate aircraft cost is to analyze aircraft mission profiles to calculate the flight duration, which is the airborne time. That said, aircraft segmental flight time is calculated via the most straightforward form which consists of a sequence of ascending → cruising → descending, based on the approach proposed by Lewe, [175] as illustrated in Figure 69. As mentioned in the above assumption, weather influence is not considered since it is assumed to express the daily representative situation.

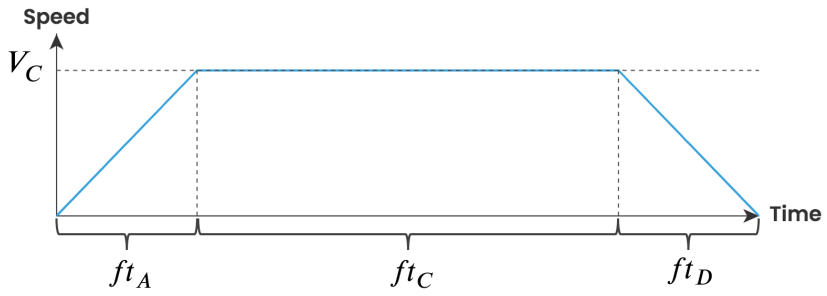


Figure 69 – Employed mission profile.

In the figure, V_C, ft_A, ft_C, ft_D represent cruising speed, ascending time, cruising time, and descending time, respectively. At last, their mathematical relationship among these parameters are expressed in the following equation:

$$ft_{i,j} = ft_A + ft_C + ft_D = ft_A + \frac{D_{i,j}}{V_C} + ft_D, \tag{B.2}$$

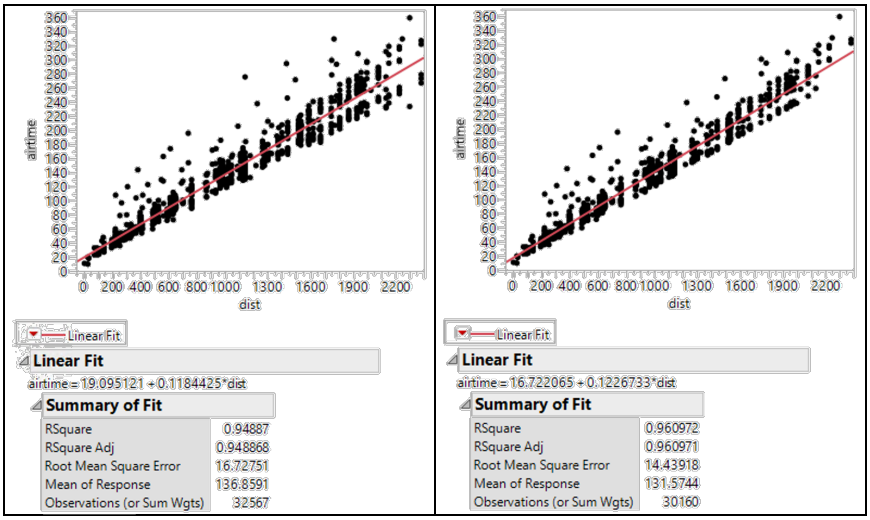
where $D_{i,j}$ is the segment distance (i.e., great-circle distance) between airport i and j.

The time for ascending and descending ($ft_{AD} = ft_A + ft_D$) is calculated by taking the intercept in the scatter plot of each aircraft type using the T-100D of all available years of 1990 to 2018. In the plot, the distance and airborne time form the abscissa and ordinate, respectively.

A pre-filtering process needs to be performed ahead because the T-100D datasets are subject to error. Some samples show a non-physically high value of the airborne time compared to the distance and vice versa. Therefore, a simple filter which confines the lower and upper bounds of the plausible airborne time is defined as expressed below:

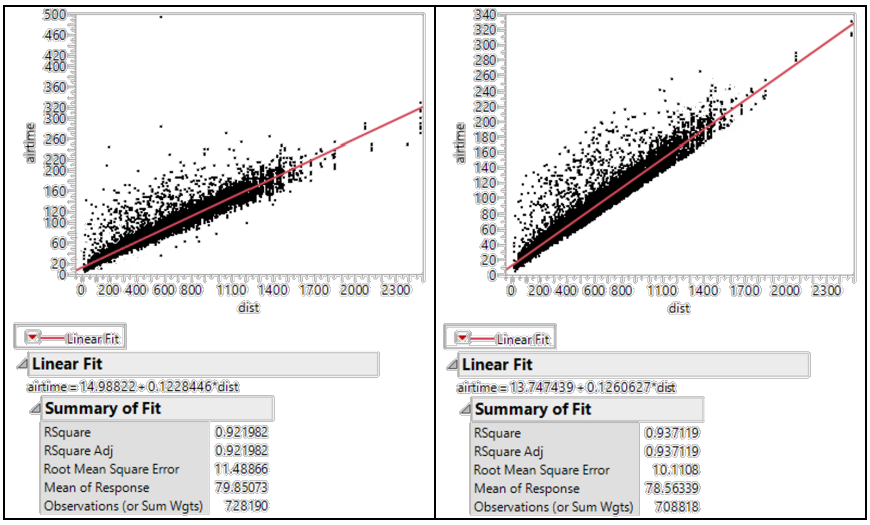
$$\frac{D_{i,j}}{V_C} < ft_{i,j} < 120 + \frac{D_{i,j}}{V_C} \tag{B.3}$$

Here, the units for time and distance are in minutes and miles. As a result, only the samples of which the airborne times are within this range are considered to estimate t_{AD} for each aircraft type. Figure 70 ~ Figure 77 illustrate the filtered linear regression for the aircraft type groups from G3 to G10 contained in the T-100D.



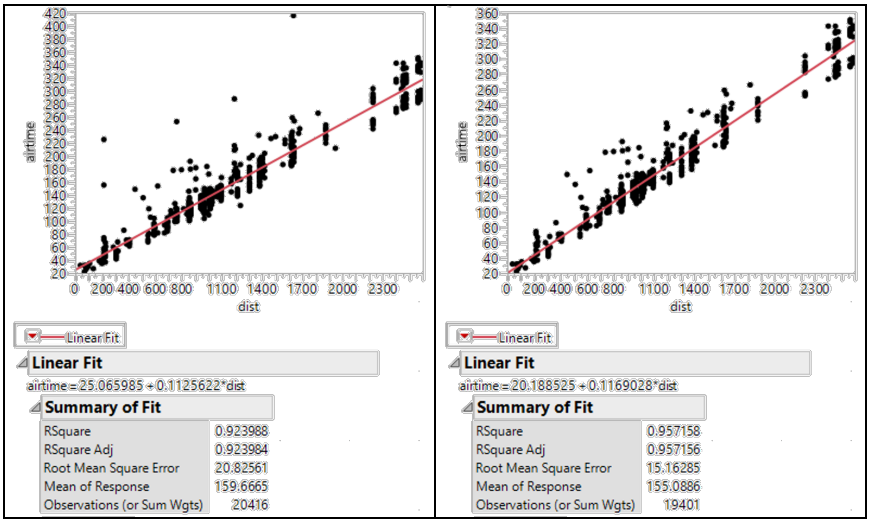
Before filtering: $ft_{AD} = 19.1$ After filtering: $ft_{AD} = 16.7$

Figure 70 – ft_{AD} regression for G3.



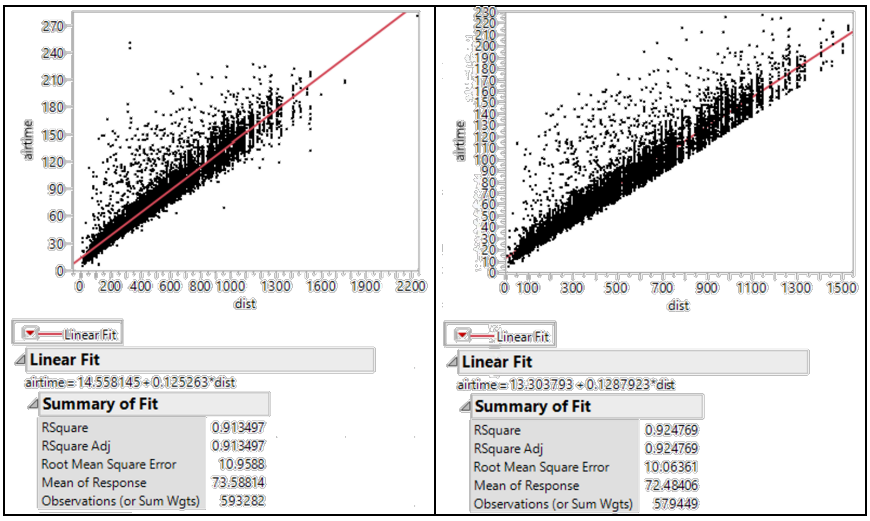
Before filtering: $ft_{AD} = 15.0$ After filtering: $ft_{AD} = 13.7$

Figure 71 – ft_{AD} regression for G4.



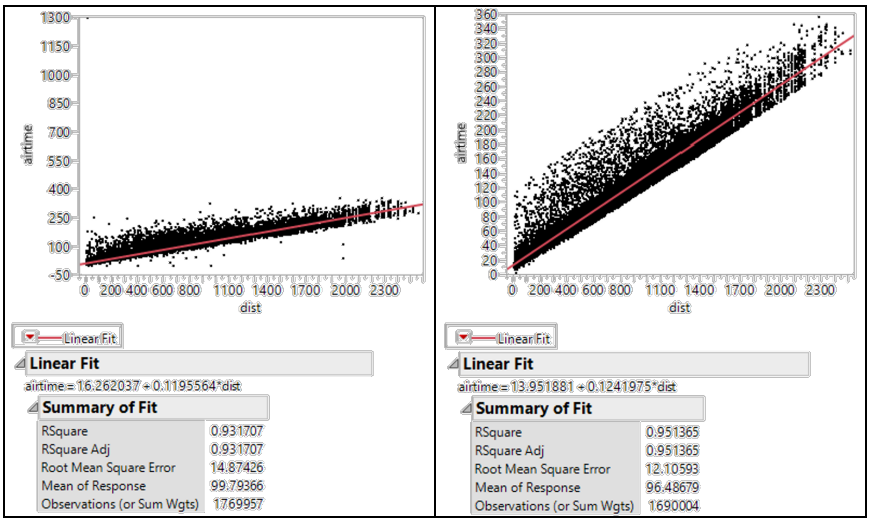
Before filtering: $ft_{AD} = 25.1$ After filtering: $ft_{AD} = 20.2$

Figure 72 – ft_{AD} regression for G5.



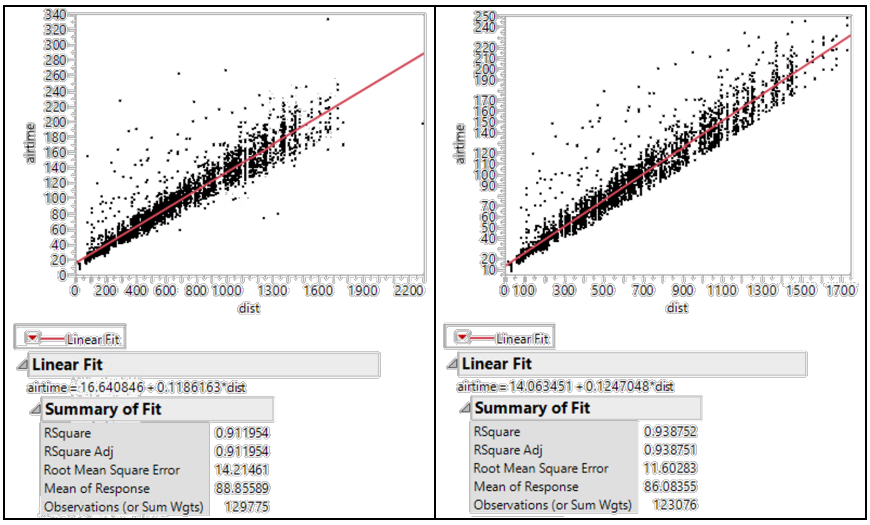
Before filtering: $ft_{AD} = 14.6$ After filtering: $ft_{AD} = 13.3$

Figure 73 – ft_{AD} regression for G6.



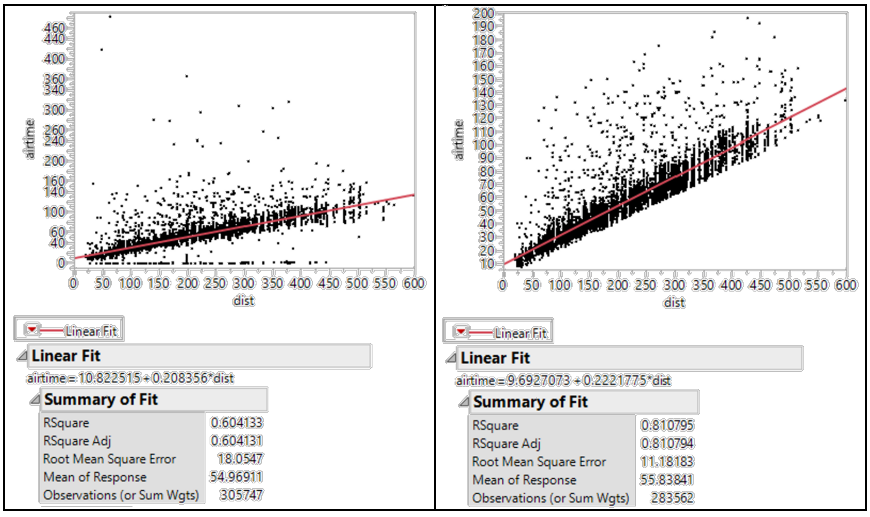
Before filtering: $ft_{AD} = 16.3$ After filtering: $ft_{AD} = 14.0$

Figure 74 – ft_{AD} regression for G7.



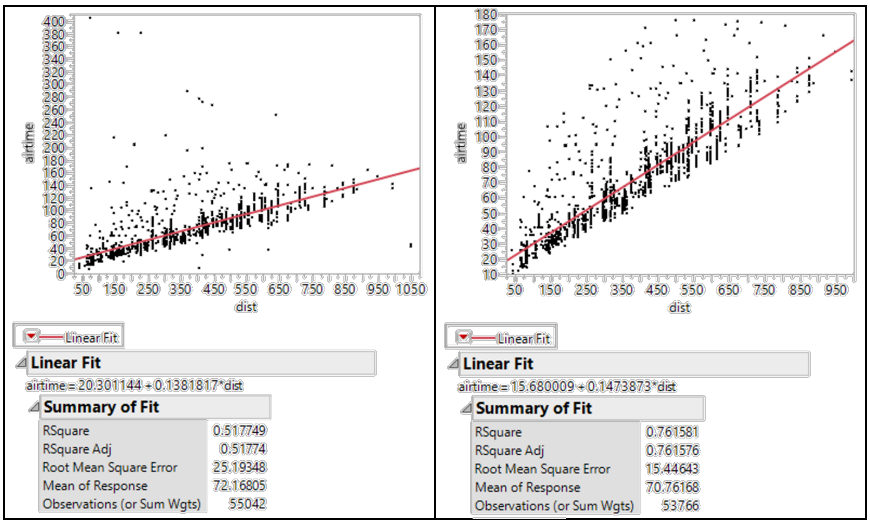
Before filtering: $ft_{AD} = 16.6$ After filtering: $ft_{AD} = 14.1$

Figure 75 – ft_{AD} regression for G8.



Before filtering: $ft_{AD} = 10.8$ After filtering: $ft_{AD} = 9.7$

Figure 76 – ft_{AD} regression for G9.



Before filtering: $ft_{AD} = 20.3$ After filtering: $ft_{AD} = 14.7$

Figure 77 – ft_{AD} regression for G10.

Note that G1 and G2 are omitted. In order to estimate their ft_{AD} values, a simple non-linear regression using a shallow neural network based on the representative specifications is conducted. The input variables are performance-related ones: capacity, range, speed. As a result, the ft_{AD} values of G1 and G2 are 9.48 minutes and 8.69 minutes, respectively. Table 40 summarizes the result of the non-linear regression, and Figure 78 illustrates the sensitivity profiles of the model. The neural network used one sigmoid function and one gaussian function.

Table 40 – Summary of fit of shallow neural regression for ft_{AD}.

Property	Training	Validation
R-square	0.9999	0.9999
RMSE	7.836E-6	2.747E-6
SSE	3.684E-10	1.509E-11

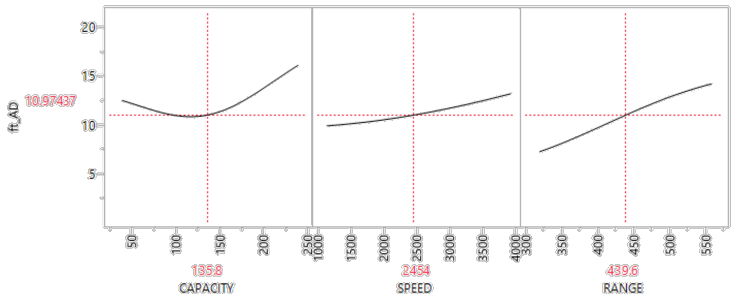


Figure 78 – Prediction profiler of ft_{AD} by neural regression for G1 & G2.

APPENDIX C. MASTER TABLE OF CONSIDERED AIRPORTS

The data in this section contains comprehensive information on the active U.S. airports in the civil ATN. Table 41 contains all 438-airport data. The data in production and attraction and demand columns are of 2018, the reference data used for verification and validation. Moreover, the values are daily representative and rounded off so that a '-' symbol indicates the demand of less than 0.5.

Table 41 – Complete airport information (2018).

Code	Latitude	Longitude	Debut Year	Production	Attraction	Demand	City	State
DAL	32.8471	-96.8518	1917	7,835	7,872	15,707	Dallas	TX
TUS	32.1161	-110.9410	1919	2,323	2,316	4,639	Tucson	AZ
MKE	42.9469	-87.8971	1920	4,761	4,821	9,582	Milwaukee	WI
BOS	42.3629	-71.0064	1923	21,281	21,401	42,682	Boston	MA
STL	38.7487	-90.3700	1923	7,655	7,713	15,368	St Louis	MO
CLE	41.4094	-81.8547	1925	6,043	6,089	12,132	Cleveland	OH
ATL	33.6367	-84.4279	1926	25,404	25,212	50,616	Atlanta	GA
MSP	44.8820	-93.2218	1926	14,758	14,737	29,495	Minneapolis	MN
BUF	42.9404	-78.7306	1926	3,317	3,353	6,670	Buffalo	NY
MFR	42.3742	-122.8735	1926	630	631	1,261	Medford	OR
BTM	45.9548	-112.4975	1926	38	37	75	Butte	MT
SFO	37.6188	-122.3754	1927	23,440	23,526	46,966	San Francisco	CA
MDW	41.7860	-87.7524	1927	9,963	10,062	20,025	Chicago	IL
HNL	21.3178	-157.9203	1927	8,267	8,188	16,455	Honolulu	HI
HOU	29.6454	-95.2789	1927	6,592	6,623	13,215	Houston	TX
RIC	37.5052	-77.3197	1927	2,559	2,556	5,115	Richmond	VA
COS	38.8058	-104.7008	1927	1,197	1,220	2,417	Colorado Springs	CO
GSO	36.1013	-79.9411	1927	1,122	1,147	2,269	Greensboro	NC
ILM	34.2711	-77.9029	1927	575	584	1,159	Wilmington	NC
PHX	33.4343	-112.0116	1928	19,942	19,829	39,771	Phoenix	AZ
EWR	40.6925	-74.1687	1928	18,126	18,224	36,350	Newark	NJ
SAN	32.7336	-117.1897	1928	15,689	15,696	31,385	San Diego	CA
TPA	27.9755	-82.5332	1928	13,249	13,191	26,440	Tampa	FL
MIA	25.7954	-80.2901	1928	10,070	10,051	20,121	Miami	FL
SYR	43.1112	-76.1063	1928	1,427	1,428	2,855	Syracuse	NY
MLB	28.1028	-80.6452	1928	315	312	627	Melbourne	FL
RST	43.9083	-92.5000	1928	219	218	437	Rochester	MN
IAG	43.1076	-78.9458	1928	179	175	354	Niagara Falls	NY
COU	38.8181	-92.2196	1928	128	140	268	Columbia	MO
FLO	34.1854	-79.7239	1928	66	64	130	Florence	SC
HGR	39.7085	-77.7265	1928	31	31	62	Hagerstown	MD
MCE	37.2848	-120.5139	1928	-	-	-	Merced	CA
FLL	26.0717	-80.1497	1929	16,709	16,626	33,335	Fort Lauderdale	FL

CMH	39.9969	-82.8922	1929	5,041	5,085	10,126	Columbus	OH
ONT	34.0560	-117.6012	1929	3,431	3,477	6,908	Ontario	CA
RNO	39.4991	-119.7681	1929	2,744	2,755	5,499	Reno	NV
TTN	40.2767	-74.8135	1929	586	576	1,162	Trenton	NJ
ABE	40.6524	-75.4404	1929	514	515	1,029	Allentown	PA
LAX	33.9426	-118.4078	1930	34,427	34,336	68,763	Los Angeles	CA
DTW	42.2124	-83.3534	1930	13,086	13,051	26,137	Detroit	MI
AUS	30.1945	-97.6699	1930	9,729	9,753	19,482	Austin	TX
CHA	35.0352	-85.2036	1930	645	648	1,293	Chattanooga	TN
ROA	37.3255	-79.9754	1930	379	384	763	Roanoke	VA
PVD	41.7240	-71.4282	1931	2,868	2,915	5,783	Providence	RI
PWM	43.6456	-70.3086	1931	1,368	1,392	2,760	Portland	ME
LIT	34.7294	-92.2248	1931	1,332	1,329	2,661	Little Rock	AR
BGR	44.8074	-68.8281	1931	395	393	788	Bangor	ME
EWN	35.0729	-77.0430	1931	151	152	303	New Bern	NC
HVN	41.2637	-72.8868	1931	51	56	107	New Haven	CT
EKO	40.8250	-115.7913	1931	21	19	40	Elko	NV
PNS	30.4734	-87.1866	1934	1,217	1,218	2,435	Pensacola	FL
SWF	41.5041	-74.1048	1934	244	252	496	New York	NY
CLT	35.2137	-80.9491	1935	9,125	8,762	17,887	Charlotte	NC
PDX	45.5887	-122.5969	1936	11,479	11,531	23,010	Portland	OR
PBI	26.6832	-80.0956	1936	4,470	4,448	8,918	West Palm Beach	FL
BOI	43.5644	-116.2229	1936	2,522	2,532	5,054	Boise	ID
MRY	36.5870	-121.8428	1936	242	247	489	Monterey	CA
BNA	36.1245	-86.6782	1937	8,960	8,964	17,924	Nashville	TN
MEM	35.0424	-89.9767	1937	2,867	2,864	5,731	Memphis	TN
CHS	32.8986	-80.0405	1937	2,867	2,863	5,730	Charleston	SC
MYR	33.6797	-78.9283	1937	1,766	1,771	3,537	Myrtle Beach	SC
MSN	43.1399	-89.3375	1937	1,296	1,305	2,601	Madison	WI
DAY	39.9022	-84.2194	1937	1,055	1,066	2,121	Dayton	OH
BTV	44.4720	-73.1533	1937	811	827	1,638	Burlington	VT
LBB	33.6637	-101.8206	1937	641	633	1,274	Lubbock	TX
FAR	46.9207	-96.8158	1937	517	517	1,034	Fargo	ND
PSC	46.2647	-119.1190	1937	474	472	946	Pasco	WA
AMA	35.2194	-101.7059	1937	458	457	915	Amarillo	TX
BIS	46.7727	-100.7458	1937	345	346	691	Bismarck	ND
TRI	36.4752	-82.4074	1937	265	260	525	Bristol/Johnson/Kingsport	TN
BRO	25.9061	-97.4260	1937	134	126	260	Brownsville	TX
LYH	37.3254	-79.2012	1937	111	113	224	Lynchburg	VA
YKM	46.5682	-120.5441	1937	87	88	175	Yakima	WA
TYR	32.3538	-95.4027	1937	66	63	129	Tyler	TX
COD	44.5202	-109.0238	1937	47	47	94	Cody	WY
ABR	45.4468	-98.4224	1937	39	37	76	Aberdeen	SD
GCC	44.3489	-105.5394	1937	30	30	60	Gillette	WY
SPN	15.1189	145.7294	1937	24	24	48	Obyan	MP
ATY	44.9140	-97.1547	1937	-	-	-	Watertown	SD
CYS	41.1556	-104.8105	1937	-	-	-	Cheyenne	WY
PDT	45.6948	-118.8430	1937	-	-	-	Pendleton	OR
RIW	43.0643	-108.4598	1937	-	-	-	Riverton	WY
SLC	40.7884	-111.9778	1938	10,076	9,970	20,046	Salt Lake City	UT
ABQ	35.0389	-106.6083	1938	3,584	3,595	7,179	Albuquerque	NM
ORF	36.8946	-76.2012	1938	2,335	2,397	4,732	Norfolk	VA
ELP	31.8073	-106.3764	1938	2,097	2,075	4,172	El Paso	TX
ITO	19.7203	-155.0485	1938	862	877	1,739	Hilo	HI
FSD	43.5820	-96.7419	1938	655	657	1,312	Sioux Falls	SD

RDM	44.2541	-121.1500	1938	518	527	1,045	Redmond	OR
BLI	48.7927	-122.5375	1938	515	516	1,031	Bellingham	WA
LBE	40.2731	-79.4103	1938	223	221	444	Latrobe	PA
LSE	43.8793	-91.2566	1938	130	124	254	La Crosse	WI
PUW	46.7439	-117.1096	1938	86	88	174	Pullman/Moscow	WA
CKB	39.2977	-80.2275	1938	36	36	72	Clarksburg	WV
IPT	41.2417	-76.9218	1938	28	28	56	Williamsport	PA
DIK	46.7973	-102.8019	1938	22	22	44	Dickinson	ND
CDC	37.7010	-113.0989	1938	20	19	39	Cedar City	UT
LAR	41.3121	-105.6750	1938	17	17	34	Laramie	WY
JMS	46.9297	-98.6782	1938	14	12	26	Jamestown	ND
YNG	41.2616	-80.6804	1938	12	12	24	Youngstown/ Warren	OH
LGA	40.7772	-73.8726	1939	18,711	18,727	37,438	New York	NY
JFK	40.6399	-73.7787	1939	15,593	15,621	31,214	New York	NY
JAC	43.6073	-110.7377	1939	454	442	896	Jackson	WY
AEX	31.3274	-92.5486	1939	135	133	268	Alexandria	LA
PSE	18.0083	-66.5630	1939	115	115	230	Ponce	PR
PHL	39.8721	-75.2407	1940	13,467	13,453	26,920	Philadelphia	PA
OAK	37.7187	-122.2217	1940	8,322	8,399	16,721	Oakland	CA
IND	39.7173	-86.2946	1940	5,972	5,995	11,967	Indianapolis	IN
BDL	41.9391	-72.6834	1940	4,407	4,466	8,873	Windsor Locks	CT
OMA	41.3032	-95.8941	1940	3,193	3,231	6,424	Omaha	NE
OKC	35.3931	-97.6008	1940	2,700	2,739	5,439	Oklahoma City	OK
LGB	33.8176	-118.1515	1940	2,695	2,714	5,409	Long Beach	CA
GEG	47.6190	-117.5352	1940	2,498	2,518	5,016	Spokane	WA
BHM	33.5639	-86.7523	1940	1,883	1,894	3,777	Birmingham	AL
DSM	41.5340	-93.6631	1940	1,695	1,716	3,411	Des Moines	IA
ROC	43.1191	-77.6719	1940	1,605	1,616	3,221	Rochester	NY
CAK	40.9151	-81.4436	1940	594	616	1,210	Akron	OH
BIL	45.8077	-108.5429	1940	537	538	1,075	Billings	MT
BTR	30.5329	-91.1499	1940	511	514	1,025	Baton Rouge	LA
SBA	34.4262	-119.8415	1940	509	504	1,013	Santa Barbara	CA
DAB	29.1799	-81.0581	1940	501	501	1,002	Daytona Beach	FL
FNT	42.9655	-83.7448	1940	496	502	998	Flint	MI
EYW	24.5561	-81.7596	1940	489	470	959	Key West	FL
MLI	41.4483	-90.5075	1940	451	454	905	Moline	IL
MFE	26.1758	-98.2386	1940	447	436	883	Mc Allen	TX
SBN	41.7082	-86.3173	1940	439	441	880	South Bend	IN
GNV	29.6901	-82.2718	1940	297	295	592	Gainesville	FL
TVC	44.7416	-85.5818	1940	294	298	592	Traverse City	MI
SBP	35.2373	-120.6426	1940	292	292	584	San Luis Obispo	CA
GJT	39.1224	-108.5267	1940	277	279	556	Grand Junction	CO
LFT	30.2050	-91.9877	1940	262	275	537	Lafayette	LA
EVV	38.0408	-87.5285	1940	256	254	510	Evansville	IN
GTF	47.4824	-111.3702	1940	221	224	445	Great Falls	MT
BMI	40.4771	-88.9159	1940	219	221	440	Bloomington/ Normal	IL
EGE	39.6427	-106.9159	1940	212	211	423	Eagle	CO
MOT	48.2576	-101.2780	1940	197	194	391	Minot	ND
LAN	42.7786	-84.5862	1940	187	197	384	Lansing	MI
ELM	42.1599	-76.8917	1940	169	172	341	Elmira/Corning	NY
DLH	46.8421	-92.1936	1940	162	168	330	Duluth	MN

AZO	42.2344	-85.5516	1940	165	163	328	Kalamazoo	MI
LNK	40.8509	-96.7591	1940	144	148	292	Lincoln	NE
SCK	37.8942	-121.2383	1940	147	145	292	Stockton	CA
HLN	46.6067	-111.9833	1940	135	140	275	Helena	MT
BFL	35.4339	-119.0577	1940	135	138	273	Bakersfield	CA
MLU	32.5109	-92.0377	1940	124	121	245	Monroe	LA
SUN	43.5038	-114.2956	1940	121	120	241	Hailey	ID
YUM	32.6567	-114.6061	1940	106	104	210	Yuma	AZ
ACK	41.2530	-70.0599	1940	90	91	181	Nantucket	MA
CLL	30.5886	-96.3638	1940	90	88	178	College Station	TX
GRI	40.9675	-98.3096	1940	83	84	167	Grand Island	NE
MHK	39.1412	-96.6718	1940	84	83	167	Manhattan	KS
VLD	30.7814	-83.2762	1940	58	58	116	Valdosta	GA
JLN	37.1518	-94.4983	1940	51	51	102	Joplin	MO
ABY	31.5355	-84.1945	1940	50	48	98	Albany	GA
TXK	33.4537	-93.9910	1940	48	48	96	Texarkana	AR
ART	43.9918	-76.0194	1940	31	35	66	Watertown	NY
ESC	45.7227	-87.0937	1940	22	19	41	Escanaba	MI
MEI	32.3326	-88.7519	1940	20	20	40	Meridian	MS
IMT	45.8184	-88.1146	1940	19	20	39	Iron Mountain Kingsford	MI
HIB	47.3866	-92.8390	1940	20	19	39	Hibbing	MN
MKG	43.1677	-86.2354	1940	18	17	35	Muskegon	MI
OTH	43.4169	-124.2470	1940	13	13	26	North Bend	OR
APN	45.0781	-83.5603	1940	11	11	22	Alpena	MI
HYA	41.6693	-70.2804	1940	9	9	18	Hyannis	MA
PRC	34.6548	-112.4192	1940	2	2	4	Prescott	AZ
PVC	42.0723	-70.2207	1940	1	2	3	Provincetown	MA
PQI	46.6890	-68.0448	1940	2	-	2	Presque Isle	ME
AUG	44.3206	-69.7973	1940	1	1	2	Augusta	ME
BHB	44.4498	-68.3616	1940	-	-	-	Bar Harbor	ME
CDR	42.8376	-103.0954	1940	-	-	-	Chadron	NE
DDC	37.7631	-99.9654	1940	-	-	-	Dodge City	KS
HVR	48.5430	-109.7623	1940	-	-	-	Havre	MT
IPL	32.8342	-115.5788	1940	-	-	-	Imperial	CA
IYK	35.6587	-117.8295	1940	-	-	-	Inyokern	CA
MSL	34.7453	-87.6102	1940	-	-	-	Muscle Shoals	AL
MSS	44.9362	-74.8451	1940	-	-	-	Massena	NY
OLF	48.0945	-105.5751	1940	-	-	-	Wolf Point	MT
MCO	28.4294	-81.3090	1941	26,451	26,474	52,925	Orlando	FL
DCA	38.8514	-77.0377	1941	13,786	13,780	27,566	Washington	DC
SNA	33.6757	-117.8682	1941	7,431	7,512	14,943	Santa Ana	CA
ALB	42.7491	-73.8020	1941	1,924	1,935	3,859	Albany	NY
AZA	33.3078	-111.6556	1941	1,084	1,081	2,165	Mesa	AZ
VPS	30.4832	-86.5260	1941	919	910	1,829	Valparaiso/Destin-Ft Walton Beach	FL
MAF	31.9425	-102.2019	1941	845	823	1,668	Midland	TX
MSO	46.9163	-114.0906	1941	535	529	1,064	Missoula	MT
FWA	40.9785	-85.1952	1941	460	460	920	Fort Wayne	IN
AGS	33.3699	-81.9645	1941	405	399	804	Augusta	GA
JNU	58.3547	-134.5785	1941	366	371	737	Juneau	AK
FSM	35.3366	-94.3674	1941	103	102	205	Fort Smith	AR
PGV	35.6357	-77.3841	1941	66	71	137	Greenville	NC
OGD	41.1957	-112.0129	1941	33	32	65	Ogden	UT

PLN	45.5709	-84.7967	1941	30	32	62	Pellston	MI
ALS	37.4351	-105.8679	1941	-	-	-	Alamosa	CO
CNM	32.3374	-104.2634	1941	-	-	-	Carlsbad	NM
IRK	40.0935	-92.5449	1941	-	-	-	Kirksville	MO
LAS	36.0801	-115.1523	1942	26,964	26,701	53,665	Las Vegas	NV
SAT	29.5340	-98.4691	1942	6,173	6,201	12,374	San Antonio	TX
PIT	40.4914	-80.2327	1942	5,888	5,953	11,841	Pittsburgh	PA
BUR	34.2007	-118.3587	1942	3,770	3,787	7,557	Burbank	CA
SDF	38.1741	-85.7365	1942	2,329	2,345	4,674	Louisville	KY
TUL	36.1984	-95.8881	1942	1,930	1,950	3,880	Tulsa	OK
SAV	32.1276	-81.2021	1942	1,725	1,743	3,468	Savannah	GA
PIE	27.9087	-82.6865	1942	1,621	1,647	3,268	St Petersburg-Clearwater	FL
PSP	33.8297	-116.5067	1942	1,261	1,259	2,520	Palm Springs	CA
FAT	36.7766	-119.7188	1942	940	951	1,891	Fresno	CA
BZN	45.7776	-111.1520	1942	858	849	1,707	Bozeman	MT
SRQ	27.3954	-82.5544	1942	851	850	1,701	Sarasota/Bradenton	FL
MDT	40.1932	-76.7626	1942	803	803	1,606	Harrisburg	PA
LEX	38.0368	-84.6086	1942	787	796	1,583	Lexington	KY
CAE	33.9388	-81.1195	1942	659	689	1,348	Columbia	SC
SGF	37.2457	-93.3886	1942	656	652	1,308	Springfield	MO
MOB	30.6914	-88.2428	1942	369	376	745	Mobile	AL
GPT	30.4073	-89.0701	1942	373	354	727	Gulfport	MS
BLV	38.5452	-89.8352	1942	231	229	460	Belleville	IL
SAF	35.6171	-106.0894	1942	136	131	267	Santa Fe	NM
PVU	40.2192	-111.7234	1942	116	114	230	Provo	UT
EAT	47.3988	-120.2068	1942	84	88	172	Wenatchee	WA
BET	60.7786	-161.8372	1942	85	85	170	Bethel	AK
ALW	46.0952	-118.2859	1942	71	73	144	Walla Walla	WA
DHN	31.3210	-85.4495	1942	65	63	128	Dothan	AL
CSG	32.5163	-84.9389	1942	63	60	123	Columbus	GA
SUX	42.4026	-96.3844	1942	50	52	102	Sioux City	IA
GCK	37.9275	-100.7244	1942	31	30	61	Garden City	KS
HOB	32.6875	-103.2173	1942	25	25	50	Hobbs	NM
RKS	41.5942	-109.0652	1942	23	22	45	Rock Springs	WY
INL	48.5656	-93.4022	1942	19	19	38	International Falls	MN
BFF	41.8740	-103.5956	1942	8	8	16	Scottsbluff	NE
CDB	55.2053	-162.7245	1942	2	2	4	Cold Bay	AK
RUT	43.5297	-72.9496	1942	1	-	1	Rutland	VT
EWB	41.6766	-70.9578	1942	-	-	-	New Bedford	MA
GLH	33.4829	-90.9856	1942	-	-	-	Greenville	MS
SLK	44.3853	-74.2062	1942	-	-	-	Saranac Lake	NY
SVC	32.6365	-108.1564	1942	-	-	-	Silver City	NM
ORD	41.9773	-87.9080	1943	26,350	26,065	52,415	Chicago	IL
RDU	35.8776	-78.7875	1943	7,680	7,712	15,392	Raleigh/Durham	NC
SFB	28.7770	-81.2349	1943	2,046	2,090	4,136	Orlando	FL
MHT	42.9328	-71.4358	1943	1,346	1,355	2,701	Manchester	NH
ISP	40.7952	-73.1002	1943	1,265	1,270	2,535	New York	NY
HPN	41.0670	-73.7076	1943	1,010	1,018	2,028	White Plains	NY
ACY	39.4576	-74.5772	1943	762	757	1,519	Atlantic City	NJ
EUG	44.1246	-123.2120	1943	722	734	1,456	Eugene	OR
HRL	26.2271	-97.6551	1943	414	404	818	Harlingen	TX
LCK	39.8138	-82.9278	1943	228	222	450	Columbus	OH
MBS	43.5329	-84.0796	1943	141	147	288	Saginaw	MI
CPR	42.9059	-106.4636	1943	113	113	226	Casper	WY

GUM	13.4840	144.7971	1943	96	92	188	Guam	GU
ACV	40.9778	-124.1085	1943	84	83	167	Arcata/Eureka	CA
ACT	31.6122	-97.2303	1943	76	77	153	Waco	TX
SJT	31.3577	-100.4963	1943	72	67	139	San Angelo	TX
DBQ	42.4020	-90.7095	1943	52	52	104	Dubuque	IA
SWO	36.1614	-97.0859	1943	29	28	57	Stillwater	OK
PAH	37.0603	-88.7730	1943	24	23	47	Paducah	KY
CGI	37.2253	-89.5707	1943	6	5	11	Cape Girardeau	MO
PIR	44.3827	-100.2860	1943	1	-	1	Pierre	SD
EAR	40.7270	-99.0068	1943	1	-	1	Kearney	NE
GGW	48.2124	-106.6148	1943	-	-	-	Glasgow	MT
CEC	41.7802	-124.2365	1943	-	-	-	Crescent City	CA
RKD	44.0601	-69.0993	1943	-	-	-	Rockland	ME
SPB	45.7710	-122.8618	1943	-	-	-	Scappoose	OR
SEA	47.4499	-122.3118	1944	21,511	21,572	43,083	Seattle	WA
CVG	39.0488	-84.6678	1944	5,269	5,257	10,526	Covington	KY
PGD	26.9189	-81.9909	1944	1,145	1,154	2,299	Punta Gorda	FL
CID	41.8847	-91.7108	1944	738	740	1,478	Cedar Rapids	IA
LWS	46.3745	-117.0154	1944	88	86	174	Lewiston	ID
SBY	38.3402	-75.5095	1944	78	84	162	Salisbury	MD
PIH	42.9098	-112.5959	1944	67	68	135	Pocatello	ID
BQK	31.2590	-81.4663	1944	53	49	102	Brunswick	GA
RDD	40.5090	-122.2934	1944	48	44	92	Redding	CA
RHI	45.6309	-89.4666	1944	34	34	68	Rhinelander	WI
BPT	29.9508	-94.0207	1944	29	30	59	Beaumont/Port Arthur	TX
EAU	44.8658	-91.4843	1944	23	22	45	Eau Claire	WI
PUB	38.2899	-104.4980	1944	6	5	11	Pueblo	CO
LBL	37.0439	-100.9600	1944	3	2	5	Liberal	KS
TBN	37.7416	-92.1407	1944	3	2	5	Fort Leonard Wood	MO
MWA	37.7550	-89.0111	1944	1	1	2	Marion	IL
AIA	42.0532	-102.8037	1944	-	-	-	Alliance	NE
CMI	40.0388	-88.2778	1945	139	139	278	Champaign/Urbana	IL
GGG	32.3840	-94.7115	1945	26	26	52	Longview	TX
VEL	40.4409	-109.5099	1945	4	3	7	Vernal	UT
DEC	39.8346	-88.8657	1945	-	-	-	Decatur	IL
TVF	48.0657	-96.1850	1945	-	-	-	Thief River Falls	MN
SJC	37.3630	-121.9286	1946	8,744	8,807	17,551	San Jose	CA
MSY	29.9933	-90.2590	1946	8,635	8,581	17,216	New Orleans	LA
OGG	20.8986	-156.4305	1946	3,995	4,017	8,012	Kahului	HI
STS	38.5097	-122.8129	1946	284	288	572	Santa Rosa	CA
DRO	37.1515	-107.7538	1946	256	252	508	Durango	CO
MTJ	38.5098	-107.8942	1946	154	152	306	Montrose	CO
RFD	42.1954	-89.0972	1946	151	148	299	Chicago/Rockford	IL
ORH	42.2671	-71.8756	1946	82	85	167	Worcester	MA
PSM	43.0779	-70.8233	1946	58	58	116	Portsmouth	NH
BJI	47.5107	-94.9347	1946	40	40	80	Bemidji	MN
LNY	20.7856	-156.9514	1946	12	13	25	Lanai City	HI
MKK	21.1529	-157.0963	1946	11	13	24	Kaunakakai	HI
IFP	35.1546	-114.5593	1946	3	2	5	Bullhead City	AZ
LEB	43.6261	-72.3042	1946	1	1	2	Lebanon	NH
MCK	40.2063	-100.5921	1946	-	-	-	Mc Cook	NE
SOW	34.2655	-110.0057	1946	-	-	-	Show Low	AZ

AVP	41.3385	-75.7234	1947	324	331	655	Wilkes-Barre/Scranton	PA
CRW	38.3759	-81.5930	1947	235	236	471	Charleston	WV
ITH	42.4914	-76.4587	1947	116	119	235	Ithaca	NY
SPI	39.8442	-89.6781	1947	101	98	199	Springfield	IL
OTZ	66.8848	-162.5981	1947	43	46	89	Kotzebue	AK
GUC	38.5343	-106.9317	1947	43	42	85	Gunnison	CO
SMX	34.8999	-120.4581	1947	33	31	64	Santa Maria	CA
MMH	37.6241	-118.8388	1947	29	33	62	Mammoth Lakes	CA
OGS	44.6822	-75.4633	1947	28	27	55	Ogdensburg	NY
ALO	42.5571	-92.4003	1947	27	26	53	Waterloo	IA
ENA	60.5733	-151.2448	1947	9	6	15	Kenai	AK
UIN	39.9430	-91.1945	1947	6	5	11	Quincy	IL
HOM	59.6450	-151.4858	1947	4	3	7	Homer	AK
ANI	61.5816	-159.5431	1947	-	-	-	Aniak	AK
FMN	36.7412	-108.2299	1947	-	-	-	Farmington	NM
MAZ	18.2557	-67.1485	1947	-	-	-	Mayaguez	PR
MCG	62.9528	-155.6071	1947	-	-	-	Mcgrath	AK
STT	18.3373	-64.9733	1948	397	400	797	Charlotte Amalie	VI
GRB	44.4846	-88.1297	1948	353	363	716	Green Bay	WI
PHF	37.1319	-76.4930	1948	271	274	545	Newport News	VA
STX	17.7015	-64.8019	1948	176	177	353	Christiansted	VI
FLG	35.1403	-111.6693	1948	96	95	191	Flagstaff	AZ
ADQ	57.7500	-152.4938	1948	75	67	142	Kodiak	AK
LAW	34.5677	-98.4166	1948	64	66	130	Lawton	OK
TWF	42.4818	-114.4877	1948	62	63	125	Twin Falls	ID
BRD	46.4042	-94.1338	1948	27	26	53	Brainerd	MN
CMX	47.1684	-88.4891	1948	24	27	51	Hancock	MI
FCA	48.3106	-114.2561	1949	402	391	793	Kalispell	MT
ASE	39.2219	-106.8682	1949	348	330	678	Aspen	CO
FAY	34.9912	-78.8803	1949	310	314	624	Fayetteville	NC
ISN	48.1779	-103.6423	1949	101	95	196	Williston	ND
OWB	37.7388	-87.1668	1949	22	22	44	Owensboro	KY
CDV	60.4918	-145.4776	1949	19	18	37	Cordova	AK
YAK	59.5033	-139.6603	1949	10	9	19	Yakutat	AK
GST	58.4253	-135.7074	1949	4	3	7	Gustavus	AK
ADK	51.8836	-176.6425	1949	3	4	7	Adak Island	AK
CEZ	37.3030	-108.6281	1949	-	-	-	Cortez	CO
BWI	39.1757	-76.6690	1950	13,100	13,160	26,260	Baltimore	MD
LIH	21.9760	-159.3390	1950	2,209	2,207	4,416	Lihue	HI
PIA	40.6642	-89.6933	1950	411	415	826	Peoria	IL
RAP	44.0453	-103.0574	1950	359	358	717	Rapid City	SD
GRK	31.0672	-97.8289	1950	175	180	355	Fort Hood/Killeen	TX
BGM	42.2084	-75.9796	1950	52	51	103	Binghamton	NY
DUT	53.8989	-166.5450	1950	22	24	46	Unalaska	AK
DLG	59.0447	-158.5055	1950	9	8	17	Dillingham	AK
GAM	63.7666	-171.7328	1950	-	-	-	Gambell	AK
HNM	20.7956	-156.0144	1950	-	-	-	Hana	HI
ANC	61.1742	-149.9982	1951	2,625	2,633	5,258	Anchorage	AK
FAI	64.8151	-147.8564	1951	626	661	1,287	Fairbanks	AK
SDY	47.7069	-104.1926	1951	-	-	-	Sidney	MT
WMO	64.6892	-163.4128	1951	-	-	-	White Mountain	AK
SHV	32.4466	-93.8256	1952	356	358	714	Shreveport	LA

HTS	38.3685	-82.5604	1952	147	146	293	Huntington	WV
ERI	42.0831	-80.1739	1952	103	102	205	Erie	PA
VDZ	61.1342	-146.2448	1952	1	1	2	Valdez	AK
DFW	32.8972	-97.0377	1953	19,408	19,165	38,573	Dallas-Fort Worth	TX
ICT	37.6499	-97.4331	1953	1,029	1,040	2,069	Wichita	KS
ABI	32.4113	-99.6819	1953	104	106	210	Abilene	TX
MUE	20.0013	-155.6681	1953	-	-	-	Kamuela	HI
AKN	58.6765	-156.6487	1954	13	12	25	King Salmon	AK
SJU	18.4394	-66.0021	1955	4,556	4,412	8,968	San Juan	PR
CHO	38.1396	-78.4523	1955	432	439	871	Charlottesville	VA
TOL	41.5868	-83.8078	1955	168	170	338	Toledo	OH
MCI	39.2976	-94.7139	1956	7,438	7,504	14,942	Kansas City	MO
OME	64.5126	-165.4444	1956	50	51	101	Nome	AK
PPG	-14.3317	-170.7115	1956	40	36	76	Pago Pago	AS
PGA	36.9261	-111.4484	1957	-	-	-	Page	AZ
SHD	38.2638	-78.8964	1958	5	5	10	Staunton/Waynesboro/Harrisonburg	VA
SDP	55.3137	-160.5214	1958	1	1	2	Sand Point	AK
IAN	66.9759	-160.4365	1958	-	-	-	Kiana	AK
IDA	43.5137	-112.0707	1959	204	200	404	Idaho Falls	ID
SCE	40.8492	-77.8486	1959	150	151	301	State College	PA
SPS	33.9888	-98.4919	1959	47	41	88	Wichita Falls	TX
CVN	34.4266	-103.0776	1959	-	-	-	Clovis	NM
CRP	27.7722	-97.5024	1960	443	442	885	Corpus Christi	TX
HYS	38.8422	-99.2732	1960	10	9	19	Hays	KS
AVL	35.4344	-82.5427	1961	749	748	1,497	Asheville	NC
TLH	30.3967	-84.3509	1961	448	452	900	Tallahassee	FL
LCH	30.1261	-93.2234	1961	63	61	124	Lake Charles	LA
IAD	38.9474	-77.4599	1962	6,232	6,236	12,468	Washington	DC
GSP	34.8957	-82.2189	1962	1,359	1,373	2,732	Greer	SC
GRR	42.8808	-85.5228	1963	2,027	2,033	4,060	Grand Rapids	MI
JAN	32.3112	-90.0759	1963	602	611	1,213	Jackson	MS
GFK	47.9473	-97.1738	1963	151	148	299	Grand Forks	ND
LBF	41.1262	-100.6837	1963	9	8	17	North Platte	NE
CNY	38.7550	-109.7548	1963	8	8	16	Moab	UT
CPX	18.3129	-65.3039	1963	-	-	-	Isla De Culebra	PR
PSG	56.8015	-132.9462	1964	24	23	47	Petersburg	AK
ROP	14.1744	145.2425	1964	-	-	-	Rota Island	MP
ATW	44.2581	-88.5191	1965	386	394	780	Appleton	WI
SLN	38.7906	-97.6522	1965	7	7	14	Salina	KS
SVA	63.6863	-170.4932	1965	-	-	-	Savoonga	AK
HDN	40.4812	-107.2177	1966	107	103	210	Hayden	CO
KKI	60.9079	-161.4351	1966	-	-	-	Akiachak	AK
PQS	61.9344	-162.8994	1966	-	-	-	Pilot Station	AK
SMF	38.6954	-121.5908	1967	7,720	7,785	15,505	Sacramento	CA
HSV	34.6372	-86.7751	1967	671	675	1,346	Huntsville	AL
ROW	33.3016	-104.5306	1967	70	67	137	Roswell	NM
HHH	32.2244	-80.6975	1967	42	49	91	Hilton Head Island	SC
JAX	30.4941	-81.6878	1968	3,991	3,982	7,973	Jacksonville	FL
VQS	18.1348	-65.4936	1968	-	-	-	Isla De Vieques	PR
IAH	29.9844	-95.3414	1969	11,578	11,550	23,128	Houston	TX
CWA	44.7776	-89.6668	1969	128	130	258	Mosinee	WI

SIT	57.0468	-135.3611	1969	96	99	195	Sitka	AK
LRD	27.5442	-99.4616	1970	103	111	214	Laredo	TX
SCC	70.1955	-148.4658	1970	53	50	103	Deadhorse	AK
STC	45.5466	-94.0599	1970	31	30	61	St Cloud	MN
LWB	37.8583	-80.3995	1970	4	5	9	Lewisburg	WV
GDV	47.1387	-104.8072	1970	-	-	-	Glendive	MT
KSM	62.0608	-163.3018	1970	-	-	-	St Mary'S	AK
KOA	19.7388	-156.0456	1971	2,061	2,054	4,115	Kailua/Kona	HI
OAJ	34.8292	-77.6121	1971	205	211	416	Jacksonville	NC
GTR	33.4483	-88.5914	1971	54	51	105	Columbus/W Point/Starkvil le	MS
KTN	55.3541	-131.7112	1973	157	155	312	Ketchikan	AK
TYS	35.8094	-83.9953	1974	1,313	1,303	2,616	Knoxville	TN
BQN	18.4949	-67.1294	1974	347	343	690	Aguadilla	PR
MVY	41.3934	-70.6139	1974	43	41	84	Vineyard Haven	MA
PIB	31.4671	-89.3371	1974	6	7	13	Hattiesburg-Laurel	MS
WYS	44.6884	-111.1176	1979	10	9	19	West Yellowstone	MT
SSB	17.7472	-64.7050	1980	-	-	-	Christiansted, St Croix	VI
RSW	26.5362	-81.7552	1983	6,325	6,261	12,586	Fort Myers	FL
TEX	37.9538	-107.9088	1986	-	-	-	Telluride	CO
MGM	32.3006	-86.3940	1987	209	208	417	Montgomery	AL
JHM	20.9629	-156.6730	1987	9	8	17	Lahaina	HI
WRG	56.4843	-132.3698	1988	14	14	28	Wrangell	AK
EMK	62.7861	-164.4908	1988	-	-	-	Emmonak	AK
SHH	66.2496	-166.0894	1988	-	-	-	Shishmaref	AK
DEN	39.8617	-104.6732	1993	27,217	27,144	54,361	Denver	CO
USA	35.3878	-80.7091	1993	208	202	410	Concord	NC
BRW	71.2849	-156.7686	1996	50	52	102	Barrow	AK
XNA	36.2816	-94.3078	1997	913	918	1,831	Fayetteville/S pringdale/	AR
MQT	46.3536	-87.3953	1999	54	50	104	Marquette	MI
SNP	57.1663	-170.2225	2001	-	1	1	St Paul Island	AK
CIU	46.2508	-84.4724	2002	24	22	46	Sault Ste Marie	MI
OOK	60.5413	-165.0872	2004	-	-	-	Toksook Bay	AK
CYF	60.1492	-164.2856	2005	-	-	-	Chefornak	AK
PBG	44.6509	-73.4681	2007	165	161	326	Plattsburgh	NY
BKG	36.5319	-93.2006	2008	11	10	21	Branson	MO
DVL	48.1166	-98.9100	2008	5	5	10	Devils Lake	ND
ECP	30.3582	-85.7956	2010	704	702	1,406	Panama City	FL
SGU	37.0364	-113.5103	2011	155	158	313	St George	UT

REFERENCES

[1] Albert, Réka, and Albert-László Barabási. "Statistical mechanics of complex networks." Reviews of modern physics 74, no. 1 (2002): 47.

[2] Onnela, J.P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J. and Barabási, A.L., 2007. Structure and tie strengths in mobile communication networks. Proceedings of the national academy of sciences, 104(18), pp.7332-7336.

[3] Choromański, Krzysztof, Michał Matuszak, and Jacek Miękisz. "Scale-free graph with preferential attachment and evolving internal vertex structure." Journal of Statistical Physics 151, no. 6 (2013): 1175-1183.

[4] Bureau of Transportation Statistics, “Airline Domestic Market Share,” https://www.transtats.bts.gov/ (Accessed October 30, 2019)

[5] Cook, Gerald N., and Jeremy Goodwin. "Airline Networks: A comparison of hub-and-spoke and point-to-point systems." Journal of Aviation/Aerospace Education & Research 17, no. 2 (2008): 1.

[6] Bureau of Transportation Statistics, “Air Carrier Statistics (Form 41 Traffic),” https://www.transtats.bts.gov/Tables.asp?DB_ID=110 (Accessed June 25, 2017)

[7] Bureau of Transportation Statistics, “Air Carrier Statistics (Form 41 Traffic),” https://www.transtats.bts.gov/Tables.asp?DB_ID=111 (Accessed June 25, 2017)

[8] Bureau of Transportation Statistics, “Air Carrier Statistics (Form 41 Traffic),” https://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=125 (Accessed June 25, 2017)

[9] Airways Magazine, “Delta Airlines route map,” https://airwaysmag.com/traveler/flying-behind-the-coconut-curtain/ (Accessed October 30, 2019)

[10] Worldwide Maps, “Great Lakes Airlines Route Map of the Great Salk Lake,” https://nebulatattoo.com/maps_rm.php (Accessed October 30, 2019)

[11] Mattos, B., and J. A. G. Fregnani. "Effects of the Airline Deregulation Act on Aeronautical Industry." Int J Adv Innovat Thoughts Ideas 3, no. 161 (2016): 2277-1891.

[12] Airline Timetables, “Southwest Airlines map, January 1997,” https://airlinetimetables.blogspot.com/2018/01/southwest-airlines-january-1997.html (Accessed October 30, 2019)

[13] Airline Maps, “Southwest Airlines map, January 1982,” https://airlinemaps.tumblr.com/post/164749970322/southwest-airlines-map-january-1982-the-southwest (Accessed October 30, 2019)

[14] Flight Connections, “Southwest Airlines routes and airport map,” https://www.flightconnections.com/route-map-southwest-airlines-wn (Accessed October 30, 2019)

[15] Bureau of Transportation Statistics, “First Quarter 2019 U.S. Airline Financial Data,” https://www.bts.gov/newsroom/first-quarter-2019-us-airline-financial-data (Accessed October 28, 2019)

[16] Bureau of Transportation Statistics, “Airline Employment Data by Month,” https://www.transtats.bts.gov/Employment/ (Accessed October 30, 2019)

[17] Federal Aviation Administration, “Air Traffic by the Numbers,” https://www.faa.gov/air_traffic/by_the_numbers/ (Accessed September 16, 2019)

[18] Federal Aviation Administration, “FAA National Forecast FY 2019-2039 Full Forecast Document and Tables,” https://www.faa.gov/data_research/aviation/aerospace_forecasts/media/FY2019-39_FAA_Aerospace_Forecast.pdf (Accessed October 30, 2019)

[19] Federal Aviation Administration, “FAA Aerospace Forecast,” https://www.faa.gov/data_research/aviation/aerospace_forecasts/media/FY2019-39_FAA_Aerospace_Forecast.pdf (Accessed March 3, 2019)

[20] Department for Transport of the United Kingdom, “Aviation 2050: the Future of UK Aviation,” https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/769695/aviation-2050-web.pdf (Accessed October 20, 2019)

[21] NASA, “NASA Aeronautics Strategic Implementation Plan 2017 Update,” https://www.nasa.gov/sites/default/files/atoms/files/sip-2017-03-23-17-high.pdf (Accessed October 1, 2019)

[22] NASA, “NASA Advisory Council Aeronautics Committee Report,” https://www.nasa.gov/sites/default/files/atoms/files/aeronautics_committee_report_dec2018_tagged.pdf (Accessed October 1, 2019)

[23] NASA, “New Aviation Horizons Initiative,” https://www.nasa.gov/sites/default/files/atoms/files/nasa-aero-10-yr-plan-508-reduced.pdf (Accessed March 3, 2017)

[24] Richwine, David, and Jay Brandon. "Quiet SuperSonic Technology (QueSST) Aircraft Preliminary Design Status and Low-Boom Flight Demonstration (LBFD) Project Update." (2018).

[25] Luongo, Cesar A., Philippe J. Masson, Taewoo Nam, Dimitri Mavris, Hyun D. Kim, Gerald V. Brown, Mark Waters, and David Hall. "Next generation more-electric aircraft: A potential application for HTS superconductors." IEEE Transactions on applied superconductivity 19, no. 3 (2009): 1055-1068.

[26] Boeing, “Time Period Reports,” http://www.boeing.com/commercial/#/orders-deliveries (Accessed October 10, 2019)

[27] Airbus, “Orders and Deliveries,” https://www.airbus.com/aircraft/market/orders-deliveries.html (Accessed October 10, 2019)

[28] Watts, D. J., and S. H. Strogatz. "Collective dynamics of “small-world” networks. nature, 393: 440–442." View Article (1998).

[29] Erdős, Paul, and Alfréd Rényi. "On the evolution of random graphs." Publ. Math. Inst. Hung. Acad. Sci 5, no. 1 (1960): 17-60.

[30] Bonnefoy, Philippe, and R. Hansman. "Scalability and evolutionary dynamics of air transportation networks in the United States." In 7th AIAA ATIO Conf, 2nd CEIAT Int'l Conf on Innov and Integr in Aero Sciences, 17th LTA Systems Tech Conf; followed by 2nd TEOS Forum, p. 7773. 2007.

[31] Bhadra, Dipasis, and Brendan Hogan. "A Preliminary Analysis of the Evolution of US Air Transportanion Network." In AIAA 5th ATIO and16th Lighter-Than-Air Sys Tech. and Balloon Systems Conferences, p. 7414. 2005.

[32] Wandelt, Sebastian, and Xiaoqian Sun. "Evolution of the international air transportation country network from 2002 to 2013." Transportation Research Part E: Logistics and Transportation Review 82 (2015): 55-78.

[33] Neal, Zachary. "The devil is in the details: Differences in air traffic networks by scale, species, and season." Social networks 38 (2014): 63-73.

[34] Newman, Mark EJ. "A measure of betweenness centrality based on random walks." Social networks 27, no. 1 (2005): 39-54.

[35] Guimera, Roger, Stefano Mossa, Adrian Turtschi, and LA Nunes Amaral. "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles." Proceedings of the National Academy of Sciences 102, no. 22 (2005): 7794-7799.

[36] DeLaurentis, Daniel, E-P. Han, and Tatsuya Kotegawa. "Network-theoretic approach for analyzing connectivity in air transportation networks." Journal of aircraft 45, no. 5 (2008): 1669-1679.

[37] Billie, J. Scott, and Rex K. Kincaid. "Analyzing Historical Changes in the Airline Transportation Network from the 1920's to Present Day." In The international multi-conference on complexity, informatics and cybernetics, Orlando, Florida. 2010.

[38] Cheung, Dorothy P., and Mehmet Hadi Gunes. "A complex network analysis of the United States air transportation." In Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), pp. 699-701. IEEE Computer Society, 2012.

[39] Gegov, Emil A., Alexander E. Gegov, Mark A. Atherton, and Fernand Gobet. "Evolution-based modelling of complex airport networks." (2011).

[40] Azzam, Mark, Uwe Klingauf, and Alexander Zock. "The accelerated growth of the worldwide air transportation network." The European Physical Journal Special Topics 215, no. 1 (2013): 35-48.

[41] Fleurquin, Pablo, José J. Ramasco, and Víctor M. Eguíluz. "Characterization of delay propagation in the US air-transportation network." Transportation journal 53, no. 3 (2014): 330-344.

[42] Wandelt, Sebastian, Xiaoqian Sun, and Jun Zhang. "Evolution of domestic airport networks: a review and comparative analysis." Transportmetrica B: Transport Dynamics 7, no. 1 (2019): 1-17.

[43] Guimerà, Roger, and Marta Sales-Pardo. "Missing and spurious interactions and the reconstruction of complex networks." Proceedings of the National Academy of Sciences 106, no. 52 (2009): 22073-22078.

[44] Barrat, Alain, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. "The architecture of complex weighted networks." Proceedings of the national academy of sciences 101, no. 11 (2004): 3747-3752.

[45] Brueckner, Jan K. "Network structure and airline scheduling." The Journal of Industrial Economics 52, no. 2 (2004): 291-312.

[46] Sales-Pardo, Marta, Roger Guimera, André A. Moreira, and Luís A. Nunes Amaral. "Extracting the hierarchical organization of complex systems." Proceedings of the National Academy of Sciences 104, no. 39 (2007): 15224-15229.

[47] Lordan, Oriol, and Jose M. Sallan. "Core and critical cities of global region airport networks." Physica A: Statistical Mechanics and its Applications 513 (2019): 724-733.

[48] Batagelj, Vladimir, and Matjaz Zaversnik. "An o (m) algorithm for cores decomposition of networks. CoRR." arXiv preprint cs.DS/0310049 37 (2003).

[49] Verma, Trivik, Nuno AM Araújo, and Hans J. Herrmann. "Revealing the structure of the world airline network." Scientific reports 4 (2014): 5638.

[50] Caves, Richard E. Air transport and its regulators: an industry study. Vol. 120. Harvard Univ Pr, 1962.

[51] Sinha, Dipendra. "The regulation and deregulation of US airlines." The Journal of Transport History 20, no. 1 (1999): 46-64.

[52] Ravich, Timothy M. "National Airline Policy." U. Miami Bus. L. Rev. 23 (2014): 1.

[53] H.R. Rep. No. 85-2360, pt. 1, at 1 (1958), reprintedin 1958 U.S.C.C.A.N. 3741.

[54] Fu, Xiaowen, Tae Hoon Oum, and Anming Zhang. "Air transport liberalization and its impacts on airline competition and air passenger traffic." Transportation Journal 49, no. 4 (2010): 24.

[55] Goetz, Andrew R., and Christopher J. Sutton. "The geography of deregulation in the US airline industry." Annals of the Association of American Geographers 87, no. 2 (1997): 238-263.

[56] Reynolds-Feighan, Aisling J. "The impact of US airline deregulation on airport traffic patterns." Geographical Analysis 30, no. 3 (1998): 234-253.

[57] Jia, Tao, Kun Qin, and Jie Shan. "An exploratory analysis on the evolution of the US airport network." Physica A: Statistical Mechanics and its Applications 413 (2014): 266-279.

[58] Lin, Jingyi, and Yifang Ban. "The evolving network structure of US airline system during 1990–2010." Physica A: Statistical Mechanics and its Applications 410 (2014): 302-312.

[59] Goetz, Andrew R., and Timothy M. Vowles. "The good, the bad, and the ugly: 30 years of US airline deregulation." Journal of Transport Geography 17, no. 4 (2009): 251-263.

[60] Bejan, Adrian, J. D. Charles, and Sylvie Lorente. "The evolution of airplanes." Journal of Applied Physics 116, no. 4 (2014): 044901.

[61] Reynolds-Feighan, Aisling. "US feeder airlines: Industry structure, networks and performance." Transportation Research Part A: Policy and Practice 117 (2018): 142-157.

[62] Brueckner, Jan K., and Vivek Pai. "Technological innovation in the airline industry: The impact of regional jets." International Journal of Industrial Organization 27, no. 1 (2009): 110-120.

[63] Azzam, Mark. "Evolution of airports from a network perspective–An analytical concept." Chinese Journal of Aeronautics 30, no. 2 (2017): 513-522.

[64] O'kelly, Morton E. "The location of interacting hub facilities." Transportation science 20, no. 2 (1986): 92-106.

[65] Taylor, Christine, and Olivier De Weck. "Coupled vehicle design and network flow optimization for air transportation systems." Journal of Aircraft 44, no. 5 (2007): 1479-1486.

[66] Campbell, James F., and Morton E. O'Kelly. "Twenty-five years of hub location research." Transportation Science 46, no. 2 (2012): 153-169.

[67] García, Sergio, Mercedes Landete, and Alfredo Marín. "New formulation and a branch-and-cut algorithm for the multiple allocation p-hub median problem." European Journal of Operational Research 220, no. 1 (2012): 48-57.

[68] Campbell, James F. "Integer programming formulations of discrete hub location problems." European Journal of Operational Research 72, no. 2 (1994): 387-405.

[69] Skorin-Kapov, Darko, Jadranka Skorin-Kapov, and Morton O'Kelly. "Tight linear programming relaxation of uncapacitated p-hub median problems." Location Science 1, no. 5 (1997): 68-69.

[70] Campbell, J. F. "Hub location problems and the p-hub median problem." Center for Business and Industrial Studies, University of Missouri, St. Louis, MO (1991).

[71] Ernst, Andreas T., and Mohan Krishnamoorthy. "Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem." European Journal of Operational Research 104, no. 1 (1998): 100-112.

[72] Aykin, Turgut. "The hub location and routing problem." European Journal of Operational Research 83, no. 1 (1995): 200-219.

[73] Brimberg, Jack, and Charles ReVelle. "A bi-objective plant location problem: cost vs. demand served." Location Science 6, no. 1-4 (1998): 121-135.

[74] da Graça Costa, Maria, Maria Eugénia Captivo, and João Clímaco. "Capacitated single allocation hub location problem—A bi-criteria approach." Computers & Operations Research 35, no. 11 (2008): 3671-3695.

[75] Shin, Kyoung Seok, Jun Hyuk Kim, and Yeo Keun Kim. "A two-leveled multi-objective symbiotic evolutionary algorithm for the hub and spoke location problem." Journal of Advanced Transportation 43, no. 4 (2009): 391-411.

[76] Makui, Ahmad, Mohammad Rostami, Ehsan Jahani, and Ahmad Nikui. "A multi-objective robust optimization model for the capacitated P-hub location problem under uncertainty." Management Science Letters 2, no. 2 (2002): 525-534.

[77] Sim, Thaddeus, Timothy J. Lowe, and Barrett W. Thomas. "The stochastic p-hub center problem with service-level constraints." Computers & Operations Research 36, no. 12 (2009): 3166-3177.

[78] Alumur, Sibel A., Stefan Nickel, and Francisco Saldanha-da-Gama. "Hub location under uncertainty." Transportation Research Part B: Methodological 46, no. 4 (2012): 529-543.

[79] Brown, John Howard. "An economic model of airline hubbing-and-spoking." Logistics and Transportation Review 27, no. 3 (1991): 225.

[80] Jaillet, Patrick, Gao Song, and Gang Yu. "Airline network design and hub location problems." Location science 4, no. 3 (1996): 195-212.

[81] Campbell, James F., Andreas T. Ernst, and Mohan Krishnamoorthy. "Hub arc location problems: part I—introduction and results." Management Science 51, no. 10 (2005): 1540-1555.

[82] Campbell, James F., Andreas T. Ernst, and Mohan Krishnamoorthy. "Hub arc location problems: part II—formulations and optimal algorithms." Management Science 51, no. 10 (2005): 1556-1571.

[83] Köksalan, Murat, and Banu Soylu. "Bicriteria p-hub location problems and evolutionary algorithms." INFORMS Journal on Computing 22, no. 4 (2010): 528-542.

[84] Mahmutogullari, Ali Irfan, and Bahar Y. Kara. "Hub location under competition." European Journal of Operational Research 250, no. 1 (2016): 214-225.

[85] Soylu, Banu, and Hatice Katip. "A multiobjective hub-airport location problem for an airline network design." European Journal of Operational Research 277, no. 2 (2019): 412-425.

[86] Yang, Ta-Hui. "Stochastic air freight hub location and flight routes planning." Applied Mathematical Modelling 33, no. 12 (2009): 4424-4430.

[87] Contreras, Ivan, Jean-François Cordeau, and Gilbert Laporte. "Stochastic uncapacitated hub location." European Journal of Operational Research 212, no. 3 (2011): 518-528.

[88] Marianov, Vladimir, and Daniel Serra. "Location models for airline hubs behaving as M/D/c queues." Computers & Operations Research 30, no. 7 (2003): 983-1003.

[89] Armacost, Andrew P., Cynthia Barnhart, Keith A. Ware, and Alysia M. Wilson. "UPS optimizes its air network." Interfaces 34, no. 1 (2004): 15-25.

[90] O'Kelly, Morton E. "Air freight hubs in the FedEx system: analysis of fuel use." Journal of Air Transport Management 36 (2014): 1-12.

[91] Potts, Renfrey B., and Robert M. Oliver. Flows in transportation networks. No. 388.1 P8. (1972).

[92] Bouchard, Richard J., and Clyde E. Pyers. "Use of gravity model for describing urban travel." Highway Research Record 88 (1965): 1-43.

[93] Teodorovi´c, D., “Airline Operations Research”, Gordon and Breach Science Publishers. (1988) ISBN: 2881246729.

[94] Kotegawa, Tatsuya, Daniel A. DeLaurentis, and Aaron Sengstacken. "Development of network restructuring models for improved air traffic forecasts." Transportation Research Part C: Emerging Technologies 18, no. 6 (2010): 937-949.

[95] Takahashi, Yuya, Rie Osawa, and Susumu Shirayama. "A Basic Study of the Forecast of Air Transportation Networks Using Different Forecasting Methods." Journal of Data Analysis and Information Processing 5, no. 02 (2017): 49.

[96] Hsu, Chaug-Ing, and Wei-Yin Eie. "Airline network design and adjustment in response to fluctuation in jet fuel prices." Mathematical and Computer Modelling 58, no. 11-12 (2013): 1791-1803.

[97] Hu, Xiao-Bing, and Ezequiel Di Paolo. "A genetic algorithm based on complex networks theory for the management of airline route networks." In Nature Inspired

Cooperative Strategies for Optimization (NICSO 2007), pp. 495-505. Springer, Berlin, Heidelberg, 2008.

[98] Teodorović, Dušan. "Multiattribute aircraft choice for airline network." Journal of Transportation Engineering 112, no. 6 (1986): 634-646.

[99] Mane, Muharram, and William A. Crossley. "Allocation and design of aircraft for on-demand air transportation with uncertain operations." Journal of aircraft 49, no. 1 (2012): 141-150.

[100] Roy, Satadru, William A. Crossley, Kenneth T. Moore, Justin S. Gray, and Joaquim Martins. "Next generation aircraft design considering airline operations and economics." In 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 1647. 2018.

[101] Taylor, Christine, and Olivier De Weck. "Coupled vehicle design and network flow optimization for air transportation systems." Journal of Aircraft 44, no. 5 (2007): 1479-1486.

[102] Yang, Ta-Hui. "A two-stage stochastic model for airline network design with uncertain demand." Transportmetrica 6, no. 3 (2010): 187-213.

[103] Davendralingam, Navindran, and William Crossley. "Concurrent aircraft design and airline network design incorporating passenger demand models." In 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO) and Aircraft Noise and Emissions Reduction Symposium (ANERS), p. 6971. 2009.

[104] Han, Lie, and Ning Zhang. "P-hub airline network design incorporating interaction between elastic demand and network structure." In Emerging Technologies for Information Systems, Computing, and Management, pp. 89-96. Springer, New York, NY, 2013.

[105] Yang, Ta-Hui. "Airline network design problem with different airport capacity constraints." Transportmetrica 4, no. 1 (2008): 33-49.

[106] Davendraingam, Navindran, and William Crossley. "Robust optimization of aircraft design and airline network design incorporating econometric trends." In 11th AIAA Aviation Technology, Integration, and Operations (ATIO)

Conference, including the AIAA Balloon Systems Conference and 19th AIAA Lighter-Than, p. 6927. 2011.

[107] Wu, Weiwei, and Songlin Zheng. "Research on Airline Network Design under Hub Airport Capacity Uncertainty." In CICTP 2014: Safe, Smart, and Sustainable Multimodal Transportation Systems, pp. 3254-3266. 2014.

[108] Mohri, Seyed Sina, Hossein Karimi, Ali Abdi Kordani, and Mahdi Nasrollahi. "Airline hub-and-spoke network design based on airport capacity envelope curve: A practical view." Computers & Industrial Engineering 125 (2018): 375-393.

[109] Kotegawa, Tatsuya. "Analyzing the evolutionary mechanisms of the air transportation system-of-systems using Network Theory and machine learning algorithms." PhD diss., Purdue University, 2012.

[110] Yang, Eunsuk. "A design methodology for evolutionary air transportation networks." PhD diss., Georgia Institute of Technology, 2009.

[111] Bureau of Transportation Statistics, “Airline Origin and Destination Survey (DB1B),” https://www.transtats.bts.gov/Tables.asp?DB_ID=125 (Accessed March 3, 2017)

[112] Federal Aviation Administration, “Passenger Boarding and All-Cargo Data for U.S. Airports,” https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/ (Accessed March 3, 2017)

[113] Airnav, http://www.airnav.com/ (Accessed May 5, 2018)

[114] Merry, John Allen. Aviation Internet Directory: A Guide to the 500 Best Web Sites. McGraw Hill Professional, 2001.

[115] Prokeraia, https://www.prokerala.com/travel/airports/ (Accessed May 5, 2018)

[116] AirportIQ5010, “Airport Master Records and Reports,” https://www.gcr1.com/5010web/ (Accessed May 5, 2018)

[117] AirportIQ5010, “Advanced Search and Export,” https://www.gcr1.com/5010WEB/advancedsearch.cfm (Accessed May 5, 2019)

[118] Airport-Data, http://www.airport-data.com/ (Accessed February 28, 2018)

[119] OpendataSoft, https://public.opendatasoft.com/explore/dataset/airports/export/?refine.type=AIRPORT (Accessed February 28, 2018)

[120] Airport Codes, “Airport Info,” https://www.air-port-codes.com/airport-info/ (Accessed February 28, 2018)

[121] Yang, Eun-Suk, Jung-Ho Lewe, and Dimitri Mavris. "Demand-centric analysis of the air transportation system." In The 26th Congress of ICAS and 8th AIAA ATIO, p. 8889. 2008.

[122] Bureau of Transportation Statistics, “Air Carriers: Master Coordinate,” https://www.transtats.bts.gov/Fields.asp (Accessed May 5, 2019)

[123] Federal Aviation Administration, “Economic Values for Investment and Regulatory Decisions,” https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost/media/ECONOMIC.pdf (Accessed May 18, 2018)

[124] The Boeing 247: the first modern airliner. 1991-12-01. ISBN 9780295970943. https://books.google.com/books?id=2zTueJhXW0QC&lpg=PP19&pg=PA192#v=onepage&q&f=false

[125] The New York Times, “Company News; Pratt Engines Are Selected for Delta’s Big Airbus Order,” https://www.nytimes.com/1992/03/10/business/company-news-pratt-engines-are-selected-for-delta-s-big-airbus-order.html (Accessed May 5, 2019)

[126] Wayback Machine, “Airbus Corporate information,” https://web.archive.org/web/20090907084244/http://www.airbus.com/en/corporate/orders_and_deliveries (Accessed March 3, 2017)

[127] Delta Flight Museum, “Aircraft by Type,” https://www.deltamuseum.org/exhibits/delta-history/aircraft-by-type/jet/ (Accessed February 28, 2018)

[128] CNN Travel, “Boeing 747 Retirement: Farewell to the ‘Queen of the Skies’,” https://www.cnn.com/travel/article/delta-boeing-747-retirement-flight/index.html (Accessed February 28, 2018)

[129] Planespotters, “Production List Search,” https://www.planespotters.net/production-list/search?fleet=American-Airlines&manufacturer=Boeing&subtype=737-300&fleetStatus=historic&refresh=1 (Accessed February 28, 2018)

[130] Airbus, “American Airlines Retires Its A300 Feet,” https://www.airbus.com/newsroom/news/en/2009/09/american-airlines-retires-its-a300-fleet.html (Accessed February 28, 2018)

[131] Airfleets, https://www.airfleets.net/home/ (Accessed February 28, 2018)

[132] Great Circle Mapper, “Featured Map for 22 March 2018,” http://www.gcmap.com/featured/20180322 (Accessed May 5, 2018)

[133] United Airlines, “Farewell, Your Majesty: United Airlines Flies the 747, the ‘Queen of the Skies’, One Last Time,” https://hub.united.com/farewell-your-majesty-united-airlines-2567373586.html (Accessed February 28, 2018)

[134] USA Today, “United Airlines Completes Its Last Boeing 747 Flight,” https://www.usatoday.com/story/travel/flights/todayinthesky/2017/11/07/united-airlines-final-boeing-747-flight-today/838922001/ (Accessed February 28, 2018)

[135] Wayback Machine, “United Airlines Swallow,” https://web.archive.org/web/20021003170259/http://www.united.com/page/article/0%2C%2C1408%2C00.html (Accessed February 28, 2018)

[136] Hartigan, John A., and Manchek A. Wong. "Algorithm AS 136: A k-means clustering algorithm." Journal of the Royal Statistical Society. Series C (Applied Statistics) 28, no. 1 (1979): 100-108.

[137] Statistical Discovery, “JMP,” https://www.jmp.com/en_us/home.html (Accessed May 5, 2019)

[138] Song, Kisun, Jung-Ho Lewe, and Dimitri N. Mavris. "A Multi-Tier Evolution Model of Air Transportation Networks." In 14th AIAA Aviation Technology, Integration, and Operations Conference, p. 3267. 2014.

[139] Lewe, Jung-Ho, Kisun Song, and Dimitri N. Mavris. "An Exploration of Evolution Paths for Air Transportation Networks." In 30th Congress of the International Council of the Aeronautical Sciences. 2016.

[140] Song, Kisun, Jung-Ho Lewe, and Dimitri N. Mavris. "Simulation Optimization for Historical Reenactment of the Air Transportation Network Evolution." In 18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, p. 3333. 2017.

[141] Song, Kisun, Jung-Ho Lewe, and Dimitri N. Mavris. "Reenacting the History of the US Air Transportation Network Evolution." In 2018 Aviation Technology, Integration, and Operations Conference, p. 4240. 2018.

[142] Gentry, Jennifer, Kent Duffy, and William J. Swedish. "Airport capacity profiles." Washington DC (2014).

[143] Bureau of Transportation Statistics, “Marketing Carrier On-Time Performance,” https://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=237 (Accessed June 12, 2019)

[144] Dijkstra, Edsger Wybe. "A note on two problems in connexion with graphs:(numerische mathematik, _1 (1959), p 269-271)." (1959).

[145] Attractive Chaos, “Programming language benchmark,” https://attractivechaos.github.io/plb/ (Accessed May 5, 2019)

[146] Java-T-Point, “Features of Node.js,” https://www.javatpoint.com/nodejs-features (Accessed May 5, 2019)

[147] British Airways, “Concorde Flying Manual,” Volume 1, 1979.

[148] Manuel d’utilization Concorde, Revision 58 du 20 Mars 2003, Volume 2

[149] Bureau Of Transportation Statistics, “T-100 International Segment (All Carriers)”, https://www.transtats.bts.gov/Fields.asp (Accessed June 17, 2019)

[150] Google Flights, https://www.google.com/flights (Accessed May 5, 2018)

[151] Boom, https://boomsupersonic.com/ (Accessed May 5, 2018)

[152] Bollobás, B. "Random Graphs Academic Press." New York (1985).

[153] Bollobás, Béla. Modern graph theory. Vol. 184. Springer Science & Business Media, 2013.

[154] D.B. West, Introduction to Graph Theory, Prentice-Hall, Englewood Cliffs, NJ, (1995).

[155] F. Harary, Graph Theory, Perseus, Cambridge, MA, (1995).

[156] Yook, Soon-Hyung, Hawoong Jeong, A-L. Barabási, and Yuhai Tu. "Weighted evolving networks." Physical review letters 86, no. 25 (2001): 5835.

[157] Onnela, J-P., Anirban Chakraborti, Kimmo Kaski, Janos Kertesz, and Antti Kanto. "Dynamics of market correlations: Taxonomy and portfolio analysis." Physical Review E 68, no. 5 (2003): 056110.

[158] Newman, M.E.J. Networks: An Introduction. Oxford, UK: Oxford University Press, (2010).

[159] Opsahl, Tore, Filip Agneessens, and John Skvoretz. "Node centrality in weighted networks: Generalizing degree and shortest paths." Social networks 32, no. 3 (2010): 245-251.

[160] Linton, C. "Freeman. A set of measures of centrality based on betweenness." Sociometry 40, no. 1 (1977): 35-41.

[161] Wasserman, Stanley, and Katherine Faust. Social network analysis: Methods and applications. Vol. 8. Cambridge university press, 1994.

[162] Scott, J. "Social Network Analysis: A Handbook. 2nd edn SAGE Publications." (2000).

[163] Yoon, Jeongah, Anselm Blumer, and Kyongbum Lee. "An algorithm for modularity analysis of directed and weighted biological networks based on edge-betweenness centrality." Bioinformatics 22, no. 24 (2006): 3106-3108.

[164] Bavelas, Alex. "Communication patterns in task-oriented groups." The Journal of the Acoustical Society of America 22, no. 6 (1950): 725-730.

[165] Sabidussi, Gert. "The centrality index of a graph." Psychometrika 31, no. 4 (1966): 581-603.

[166] Batagelj, Vladimir, and Andrej Mrvar. "A subquadratic triad census algorithm for large sparse networks with small maximum degree." Social networks 23, no. 3 (2001): 237-243.

[167] Davis, James A., and Samuel Leinhardt. "The structure of positive interpersonal relations in small groups." (1967).

[168] Faust, Katherine. "Comparing social networks: size, density, and local structure." Metodoloski zvezki 3, no. 2 (2006): 185.

[169] Holland, P.W. and Leinhardt, S. “A method for detecting structure in sociometric data.” American Journal of Sociology, 76, 492-513 (1970).

[170] Holland, P.W. and Leinhardt, S. “Transitivity in structural models of small groups.” Comparative Group Studies, 2, 107-124 (1971).

[171] Holland, P.W. and Leinhardt, S. “Some evidence on the transitivity of positive interpersonal sentiment.” American Journal of Sociology, 77, 1205- 1209 (1972).

[172] Macdonald, P. J., E. Almaas, and A-L. Barabási. "Minimum spanning trees of weighted scale-free networks." EPL (Europhysics Letters) 72, no. 2 (2005): 308.

[173] Jung, Woo-Sung, Fengzhong Wang, and H. Eugene Stanley. "Gravity model in the Korean highway." EPL (Europhysics Letters) 81, no. 4 (2008): 48005.

[174] Masuda, Naoki, Hiroyoshi Miwa, and Norio Konno. "Geographical threshold graphs with small-world and scale-free properties." Physical Review E 71, no. 3 (2005): 036108.

[175] Lewe, Jung-Ho, “An Integrated Decision-Making Framework for Transportation Architectures: Application to Aviation Systems Design.” PhD thesis, Georgia Institute of Technology, (2005).